

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2001, 20:20:38 : Search time 177.69 Seconds
(without alignments)
8841.296 Million cell updates/sec

Title: US-09-578-827A-4

Perfect score: 2502
Sequence: 1 agagaagcagcgctggggttt.....ttcattaaaaaaaaaatg 2502

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: N_Geneseq_0601.*
2: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn/NA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	203.4	8.1	936	AAF58252	Oligonucleotide D1
2	203.4	8.1	936	AAF58254	Oligonucleotide D1
3	203.4	8.1	936	AAF58257	Oligonucleotide D1
4	203.4	8.1	936	AAF58259	Oligonucleotide D2
5	203.4	8.1	936	AAF58262	Oligonucleotide D2
6	203.4	8.1	936	AAF58265	Oligonucleotide D1
7	199.4	7.8	936	AAF58262	Oligonucleotide D1
8	199.4	7.8	936	AAF58264	Oligonucleotide D1
9	195.4	7.8	936	AAF58257	Oligonucleotide D1
10	195.4	7.8	936	AAF58259	Oligonucleotide D2
11	195.4	7.8	936	AAF58262	Oligonucleotide D2

C 12	195.4	7.8	938	22	AAF58255	Oligonucleotide D1
C 13	68.2	2.7	244	22	AAF58238	Oligonucleotide D1
14	67.6	2.7	244	22	AAF58238	Oligonucleotide D1
15	58.4	2.3	664	20	AA33181	Base sequence of t
16	58.4	2.3	7372	20	AA33182	Base sequence of t
17	58.4	2.3	7797	20	AA33180	Coxpox virus bar f
18	58.4	2.3	7996	20	AA33184	Base sequence of t
19	56.2	2.2	2418	13	AA272886	P.falciparum GBP3
20	56	2.2	19124	18	AA722882	Plasmodium var-7 g
C 21	56	2.2	19124	21	AA98287	Plasmodium var-7 g
C 22	51.4	2.1	605	17	AA31530	Human 3' apolipop
C 23	49.8	2.0	665	21	AA31996	Human apolipoprote
C 24	48.8	2.0	19124	18	AA722882	Plasmodium var-7 g
25	48.8	2.0	19124	21	AA98287	Plasmodium var-7 g
C 26	48.6	1.9	549	21	AA94393	Cat flea head and
C 27	48.6	1.9	549	21	AA95142	Cat flea head and
C 28	47.8	1.9	3101	11	AA02047	Sequence encoding
29	47.8	1.9	20674	21	AA58017	Archidonic acid m
30	47.6	1.9	1052	10	AA90224	Malari-specific D
31	47.2	1.9	8920	15	AA62924	Cardamyl-Phosphat
32	46.8	1.9	1212	21	AA70218	Plasmodium falcipa
C 33	46.8	1.9	1711	19	AAV3336	Plasmodium berghei
C 34	46.8	1.9	4590	7	AA60472	Sequence encoding
C 35	46.8	1.9	9048	18	AA743225	Brassica napus RCA
C 36	46.4	1.9	5059	20	AA84332	Stealth virus nucl
C 37	46.2	1.8	2040	21	AA70158	Plasmodium falcipa
C 38	46.2	1.8	5760	6	AA50530	Sequence encoding
C 39	46	1.8	6621	21	AA70188	Plasmodium falcipa
40	45.8	1.8	11011	21	AA68252	B. burgdorferi tox
41	45.8	1.8	910715	20	AA20248	Borrelia burgdorfe
42	45.4	1.8	5852	12	AA011710	Dicystostellum plas
43	45.4	1.8	977	21	AA26706	Candida albicans p
44	45	1.8	1034	21	AA52527	Human secreted pro
45	45	1.8	3101	11	AA02047	Sequence encoding

ALIGNMENTS

```

RESULT 1
AAF58252
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252:
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN NC020107665-A2.
XX
PR 17-MAR-2000; 2000US-0190259.
XX
PF 26-JUL-2000; 2000MO-US20476.
XX
PD 01-FEB-2001.
XX
PI Umek RM;
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group; useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX

```

PS Example 6; Page 127; 159pp; English.

CC The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.

XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 8.1%; Score 203.4; DB 22; Length 936;
 Best Local Similarity 3.9%; Pred. NO. 6.3e-29;
 Matches 33; Conservative 551; Mismatches 267; Indels 0; Gaps 0;

```

QY 1504 tactgaacacttggtgttcacgtatttggcctatatataaacaacatgtaatt 1563
DB 1 .....
QY 1564 atatacgaatttttcggaattttagcgcatactgtaagatataacatgcatgac 1623
DB 61 .....
QY 1624 gtcttcaaatcatatgatagaacgatacgaagtagtactactcctcaaatattgcat 1683
DB 121 .....
QY 1684 gagagagatatgtattatataatttatttttgaagaagaataagagggaagtactcg 1743
DB 181 .....
QY 1744 ggtgagtcgatgtaaaacaaagaagaagaacgaaccataacgcatcatgata 1803
DB 241 .....
QY 1804 tcgacctcttactcttctccttatttatttctcagagacttttctactaat 1863
DB 301 .....
QY 1864 gaacactccaactatctaactatacaccatcgtatagaataagaataatataaga 1923
DB 361 .....
QY 1924 tactgtgatatttttaactagaataatattgctcgttaattttcgttaagtaaat 1983
DB 421 .....
QY 1984 caacattttcagtagaacaataatctacgcaaaagtagatcatttltgtccaa 2043
DB 481 .....
QY 2044 aatctcaagtagatagtggttagtaaaaaaacaacattctgattggcccaaaa 2103
DB 541 .....
QY 2104 aataaagagagaagaataatgttcaaaagtgctctctctctaaattgttttc 2163
DB 601 .....
QY 2164 actaaacccaatagattcaaacagctctacaagtcctcaaaagaataacatggagacaacaa 2223
DB 661 .....
QY 2224 ttgatgcaaaaatactctcttcatagctcttttttatctctagctttaaattact 2283
DB 721 .....
QY 2284 aataaactcacaaatccacaacccattctctacaactcacctcatcattagatttac 2343
DB 781 .....
QY 2344 cgaactccacac 2354
  
```

DB 841 cccccccccc 851

RESULT 2

ID AAF58254 standard; DNA; 936 BP.

XX AAF58254;

DT 24-APR-2001 (first entry)

DE oligonucleotide D1875.

KW electron-transfer group; ETM; mismatch; genotyping;

XX gene expression; ss.

OS Synthetic.

PN WO200107665-A2.

PD 01-FEB-2001.

PF 26-JUL-2000; 2000MO-US20476.

PR 26-JUL-1999; 9905-0145695.

PR 17-MAR-2000; 2000US-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

XX umek RM;

DR WPI; 2001-159728/16.

PT Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface

PS Example 6; Page 127; 159pp; English.

CC The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.

SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 8.1%; Score 203.4; DB 22; Length 936;
 Best Local Similarity 3.9%; Pred. NO. 6.3e-29;
 Matches 33; Conservative 551; Mismatches 267; Indels 0; Gaps 0;

```

QY 1504 tactgaacacttggtgttcacgtatttggcctatatataaacaacatgtaatt 1563
DB 1 .....
QY 1564 atatacgaatttttcggaattttagcgcatactgtaagatataacatgcatgac 1623
DB 61 .....
QY 1624 gtcttcaaatcatatgatagaacgatacgaagtagtactactcctcaaatattgcat 1683
DB 121 .....
QY 1684 gagagagatatgtattatataatttatttttgaagaagaataagagggaagtactcg 1743
DB 181 .....
QY 1744 ggtgagtcgatgtaaaacaaagaagaagaacgaaccataacgcatcatgata 1803
DB 241 .....
  
```

```

OY 1804 tcgaccttctatcttctctcttatttatttctcaggacttttctactaat 1863
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 360
OY 1864 gaaacctccaacttcttaactaatacctccatgtagaataaagaataataaga 1923
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 361 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 420
OY 1924 tatgtgatatttgaactagaataatatttgcctcgttaatttcgtaagttaat 1983
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 421 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 480
OY 1984 caacatttccagtagaagaacaaatattctgcaaaagtagatcatttctgcca 2043
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 481 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 540
OY 2044 aatccagttagctataggtgttagtaaaacaaacacatttgcctcccaaa 2103
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 541 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 600
OY 2104 aataagagagaagaatatgttcaaaagtgctctctctcctaattagtctc 2163
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 601 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 660
OY 2164 actaaacccaattagatcaaaacagctcaaaagtcacaaagataacatggacaaca 2223
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 661 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 720
OY 2224 ttcatgtcaaaaaaccccttctcctctcttcttctcctcagctctttaaatact 2283
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 721 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 780
OY 2284 aataaaactcaacaatccacaaacccatctctctacaaatccatccatctagattac 2343
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 781 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 840
OY 2344 ccaactccacac 2354
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 841 cccccccccc 851

```

RESULT 3

```

AF58257
ID AF58257 standard; DNA; 936 BP.
AC AF58257;
XX
XX 24-APR-2001 (first entry)
DE
DE Oligonucleotide D1954.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
XX Synthetic.
OS
XX MO200107665-A2.
XX
XX 01-FEB-2001.
PD
XX 26-JUL-2000; 2000WO-US20476.
PF
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT

```

```

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
PS
PS Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;
SQ

```

```

Query Match 8.1%; Score 203.4; DB 22; Length 936;
Best Local Similarity 3.9%; Pred. No. 6.3e-29;
Matches 33; Conservative 551; Mismatches 267; Indels 0; Caps 0;

```

```

OY 1504 tactgaacacctgggttccatgatttgcctcatataaacaacatcgttaatt 1563
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 60
OY 1564 atatacggatttttctcggaatttaacgcacatctgtaaglatataaacatgcatgic 1623
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 120
OY 1624 gtttcaaatcatatgatgaacatccagtaagtgtctactactcctacaattgcat 1683
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 180
OY 1684 gagaagatatgtatttaaatatttgaagaagaataagagggaagttaacttg 1743
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 240
OY 1744 ggtgcatcgatgtaaaacaaagaagaagaacccaactaagccatatacatgata 1803
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 241 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 300
OY 1804 tcgaccttctatcttctctcttatttatttctcagacttttctactaat 1863
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 360
OY 1864 gaaacctccaactataactaatacctccatgtagaataaagaataataaga 1923
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 361 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 420
OY 1924 tatgtgatatttgaactagaataatatttgcctcgttaatttcgtaagttaat 1983
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 421 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 480
OY 1984 caacatttccagtagaagaacaaatattctgcaaaagtagatcatttctgcca 2043
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 481 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 540
OY 2044 aatccagttagctataggtgttagtaaaacaaacacatttgcctcccaaa 2103
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 541 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 600
OY 2104 aataagagagaagaatatttctcaaaagtgctctctcctaatttctgcttc 2163
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 601 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 660
OY 2164 actaaacccaattagatcaaaacagctcaaaagtcacaaagataacatggacaaca 2223
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 661 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 720
OY 2224 ttcatgtcaaaaaaccccttctcctcttcttctcctcagctttaaatact 2283
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 721 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 780
OY 2284 aataaaactcaacaatccacaaacccatctctacaaatccatccatctagattac 2343
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


Db	181	www.....	240
OY	1744	ggtggatcgcgatgtgaacaacaaagaagaaaagcgaaaccctaaagccattactagtta	18033
Db	241	300
OY	1804	tgcacctcttatcttttccctcttatttatttttcacgaacttttctactaat	18633
Db	301	360
OY	1864	gaaacctccaactactcaactaacactcccatgttagataaagaatatataaga	19233
Db	361	420
OY	1924	tatgttgatatatttgractagaaaaatatattgtctctgtaattttcgtaaagtaaat	19833
Db	421	480
OY	1984	caacatttttcagtagaacaacaaatattactgcacaaaaglagalcatlattttgtccaa	20433
Db	481	540
OY	2044	aatctcgttcagtcataagggttgtagtaaaacaaacacattcttgattgtccccaaa	21033
Db	541	600
OY	2104	aataagagagagaagaatatgttccaagaagtgtctctctctctaataatgttttc	21633
Db	601	660
OY	2164	actaaacccaattgattcataacacagctctacaagctccaaaagataaacatgygacaaca	22233
Db	661	720
OY	2224	ttcgatgcacaaaatctctcttcttcacgtctctttttattctctagtctttaaataact	22833
Db	721	780
OY	2284	aataanaactcaacaatcacacaaaccattctctacaactcaactcactagatttac	23433
Db	781	840
OY	2344	ccactcccacc 2354	
Db	841	cccccctccc 851	
RESULT 7			
ID	AAF58252/c		
XX AC	AAF58252 standard; DNA; 936 BP.		
XX AC	AAF58252;		
XX DT	24-APR-2001 (first entry)		
XX DE	Oligonucleotide D1835.		
XX KM	Electron-transfer group; ETM; mismatch; genotyping;		
XX KM	gene expression; ss.		
OS	Synthetic.		
PN	MO200107665-AZ.		
PD	01-FEB-2001.		
Pf	26-JUL-2000; 2000MO-US20476.		
PR	26-JUL-1999; 99US-0145695.		
PR	17-MAR-2000; 2000US-0190259.		
PA	(CLIN-) CLINICAL MICRO SENSORS INC.		

[illegible]

QY 2273 tttaattactaatataaaa 2291

[illegible][illegible]

RESULT	11
AAF58262/C	
ID	AAF58262 standard; DNA; 936 BP.
XX	
AC	AAF58262;
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Oligonucleotide D2007.
XX	
KW	Electron-transfer group; ETW; mismatch; genotyping;
KW	gene expression; ss.
XX	
OS	Synthetic.
XX	
PN	W0200107665-A2.
XX	
PD	01-FEB-2001.
XX	
PE	26-JUL-2000; 2000MO-US20476.
XX	
PR	26-JUL-1999; 99US-0145695.
PR	17-MAR-2000; 2000US-0190259.
XX	
PA	(CLIN-) CLINICAL MICRO SENSORS INC.
XX	
PI	Unex RM;
XX	
DR	WPI; 2001-159728/16.
PT	
PT	Nucleic acids containing electron-transfer group, useful as labels in
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT	a single surface -
PS	Example 6, Page 128; 159pp; English.

```

Query Match      7.88; Score 195.4; DB 22; Length 936;
Best Local Similarity 1.38; Pred. No. 1.9e-27;
Matches 10; Conservative 549; Mismatches 240; Indels 0; Gaps 0;

```

RESULT 12

AAF58255/c
ID AAF58255 standard; DNA; 938 BP.
XX
AC AAF58255;
XX
XX 24-APR-2001 (first entry)
DE Oligonucleotide D1876.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
XX
XX Synthetic.
OS
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
XX 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match 7.8%; Score 195.4; DB 22; Length 938;
Best Local Similarity 1.3%; Pred. No. 1.9e-27;
Matches 10; Conservative 549; Mismatches 240; Indels 0; Gaps 0;

QY 1493 gggatcgatgactgaaacacttggtggttcattgattgacctatataatacaaa 1552
DB 801 GGG 742
QY 1553 acatcgtaatatatacgatttttcggaatttaagcatalctgtgaatatata 1612
DB 741 WW 682
QY 1613 acatgcatgctgttcaaatcatatgatgaacgatccaagtaagtgcactacccta 1672
DB 681 WW 622
QY 1673 caatattcatgagagatatgtattataattttgaagaagaataaagag 1732
DB 621 WW 562
QY 1733 aaggttaacttggtgcatcgatctgaacaaagaagaagaagcgaaaccctaagcc 1792
DB 561 WW 502
QY 1793 attaatgatatgacctcttattcttcctctatttttttcacagacttt 1852
DB 501 WW 442

QY 1853 ttctacttaatgaaacctccaaactataacttaacttaactccatgtagaataaaga 1912
DB 441 WW 382
QY 1913 attataaagaatgctgtgataattgtcaactagaataatataattgctctgtaatttc 1972
DB 381 WW 322
QY 1973 graagtaaatcaaatcttttcagtagaacaacattacttcgaaagtagatcat 2032
DB 321 WW 262
QY 2033 ttctgtccaaatctcagttagctataggtttagttaaacaacaaacacatttcat 2092
DB 261 WW 202
QY 2093 ttgcccacaaataaagaagaagaataatgttcaaaagtgctctctctcta 2152
DB 201 WW 142
QY 2153 attatgtttcactaaaccatagatcaacagctctacaagtcocaaagataaaca 2212
DB 141 WW 82
QY 2213 tgggaacaacattcgatgcaaaaaatccctcttttcattgctctttttatctcagct 2272
DB 81 WW 22
QY 2273 tttaattactataaaa 2291
DB 21 WW 3

RESULT 13
AAF58238/c
ID AAF58238 standard; DNA; 244 BP.
XX
AC AAF58238;
XX
XX 24-APR-2001 (first entry)
DE Oligonucleotide D1250:D1102.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
XX
OS Synthetic.
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
XX 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 4; Page 120; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)

required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the 1- α -25(OH) $_2$ D $_3$ gene.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2001, 13:44:08 ; Search time 92.89 Seconds
(without alignments)
6100.204 Million cell updates/sec

Title: US-09-578-827A-4

Perfect score: 2502

Sequence: 1 agaacgagcgcgtggggtt.....ttcataaaaaaaaaaatg 2502

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 segs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
C 1	96.2	3.8	7218	1	US-08-232-463-14 Sequence 14, Appl
C 2	56	2.2	19124	2	US-08-487-8268-13 Sequence 13, Appl
C 3	49.8	2.0	665	2	US-08-883-795A-36 Sequence 36, Appl
C 4	48.8	2.0	19124	2	US-08-487-8268-13 Sequence 13, Appl
C 5	47.2	1.9	7218	1	US-08-232-463-14 Sequence 14, Appl
C 6	47.2	1.9	8920	2	US-08-446-855A-1 Sequence 1, Appl
C 7	47.2	1.9	8920	4	US-09-150-741-1 Sequence 1, Appl
C 8	46.8	1.9	289	4	US-09-007-003-17 Sequence 17, Appl
C 9	46.8	1.9	289	4	US-09-244-796-17 Sequence 17, Appl
C 10	46.8	1.9	9048	3	US-08-973-273-4 Sequence 4, Appl
C 11	44	1.8	827	4	US-08-998-416-535 Sequence 535, App
C 12	43.6	1.7	782	4	US-08-989-478-1 Sequence 224, App
C 13	43.4	1.7	5655	3	US-08-996-685-1 Sequence 1, Appl
C 14	43.4	1.7	5655	3	US-08-996-685-1 Sequence 1, Appl
C 15	43.4	1.7	6243	3	US-09-056-075-1 Sequence 2, Appl
C 16	43.4	1.7	6243	3	US-09-056-075-1 Sequence 2, Appl
C 17	43.4	1.7	8457	1	US-08-107-755A-1 Sequence 1, Appl
C 18	43.4	1.7	8457	1	US-08-107-755A-1 Sequence 1, Appl
C 19	43.4	1.7	8457	1	US-08-544-332-1 Sequence 1, Appl
C 20	43.4	1.7	9919	3	US-08-880-179-1 Sequence 1, Appl
C 21	43.2	1.7	7244	4	US-08-378-313-26 Sequence 26, Appl
C 22	42.2	1.7	2454	1	US-09-419-459-1 Sequence 3, Appl
C 23	41.4	1.7	4507	2	US-08-568-459A-3 Sequence 3, Appl
C 24	41.4	1.7	4507	2	US-08-568-459A-3 Sequence 3, Appl
C 25	41.4	1.7	6060	5	PCT-US96-09430-7 Sequence 7, Appl
C 26	41.4	1.7	6152	4	US-08-973-462-1 Sequence 1, Appl
C 27	41.4	1.7	6152	4	US-08-973-462-1 Sequence 1, Appl

C 28	41.4	1.7	6768	1	US-08-107-755A-1 Sequence 1, Appl
C 29	41.4	1.7	8457	1	US-07-991-867B-1 Sequence 1, Appl
C 30	41.4	1.7	8457	2	US-08-544-332-1 Sequence 1, Appl
C 31	41.2	1.6	615	4	US-08-998-416-186 Sequence 186, App
C 32	41.2	1.6	773	4	US-08-998-416-385 Sequence 385, App
C 33	40.8	1.6	834	4	US-08-998-416-305 Sequence 305, App
C 34	40.8	1.6	5852	1	US-07-867-106-2 Sequence 595, App
C 35	40.6	1.6	660	4	US-07-991-867B-32 Sequence 32, Appl
C 36	40.4	1.6	658	1	US-08-107-755A-8 Sequence 8, Appl
C 37	40.4	1.6	660	2	US-08-544-332-32 Sequence 32, Appl
C 38	40.4	1.6	660	2	US-08-544-332-32 Sequence 32, Appl
C 39	40.4	1.6	1511	1	US-07-991-867B-8 Sequence 8, Appl
C 40	40.4	1.6	1511	1	US-08-107-755A-8 Sequence 8, Appl
C 41	40.4	1.6	1511	2	US-08-544-332-8 Sequence 8, Appl
C 42	40.4	1.6	4810	3	US-08-852-629-11 Sequence 11, Appl
C 43	40.4	1.6	4838	3	US-08-852-629-15 Sequence 15, Appl
C 44	40.4	1.6	5361	4	US-08-973-462-2 Sequence 2, Appl
C 45	40.2	1.6	636	4	US-08-998-416-1137 Sequence 1137, Ap

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasmidogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: Rh 32
US-08-883-795A-36

Query Match 2.0%; Score 49.8; DB 2; Length 665;
Best Local Similarity 49.2%; Pred. No. 0.028;
Matches 160; Conservative 0; Mismatches 162; Indels 3; Gaps 1;
QY 995 tatgaacattctctctattataatataatcacatgctctatgtaactgtgtaaa 1054
DB 343 TATGAACATTCTCTCTATTATAATATGTAATTAATTAACATTTTATTAATAAATTTTAA 284
QY 1055 tggtagaccacttaagtttatataatcatgatatatcttataggttcatacaaaatgg 1114
DB 283 TATGAACATTCTCTCTATTATAATTAATTAATTAATAAATTTTAA--ATTATTAATAATTTT 227
QY 1115 tcatgaacattctgaacttcaactctactctgctcattgtagatgctgactttcacatgt 1174
DB 226 AATTATTAATAATTTTAAATTAATAAATTAATTAATAATTTTAAATTAATAAATTTT 167
QY 1175 ttgaaacattagctcgatcgatcgaaactcttaacttaagcatgtttgctggtcaagctt 1234
DB 166 AATTATTAATAATTTTAAATTAATAAATTTTAAATTAATAAATTTTAAATTAATAAATTTT 107
QY 1235 aattcttgattatgtagtcaaaatcagaagcgttcagaaacttaacttaacttctc 1294
DB 106 AATTATTAATAATTTTAAATTAATAAATTTTAAATTAATAAATTTTAAATTAATAAATTTT 47
QY 1295 aaaaatacagatcaagaagaataa 1319
DB 46 AATTATTAATAATTTTAAATTAATAA 22

RESULT 4
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chlunis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhuan
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-8550
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 2.0%; Score 48.8; DB 2; Length 19124;
Best Local Similarity 43.3%; Pred. No. 0.1;
Matches 405; Conservative 0; Mismatches 512; Indels 19; Gaps 3;
QY 1564 atatacgaattttttcggaattttacgcacatctgtaagtabatataaacatgcatgctc 1623
DB 15301 ATATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15360
QY 1624 gtttcaaatcatatagatgaagatccagtgatgctactactccttaaatatgcat 1683
DB 15361 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15420
QY 1684 gagagagatgtaattatataatatttttgaagaagaagaagaagaagaagaagaagaaga 1743
DB 15421 AAAAAAATTAATAAATTTTAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAA 15480
QY 1744 ggtgagatcgatgtaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1803
DB 15481 AATTATTAATAAATTTTAAATTAATAAATTTTAAATTAATAAATTTTAAATTAATAAATTTT 15540
QY 1804 tcgacctcttactcttct 1863
DB 15541 AAAATTTTAAATTAATAAATTTTAAATTAATAAATTTTAAATTAATAAATTTTAAATTAATA 15600
QY 1864 gaaaccccaaatctataactaataacacccatccatgtagaataaagaagaagaagaagaaga 1923
DB 15601 AATTATTAATAAATTTTAAATTAATAAATTTTAAATTAATAAATTTTAAATTAATAAATTTTAA 15660

```

1 NAME: BENT, Stephen A.
2 REGISTRATION NUMBER: 29,768
3 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: (703)836-9300
6 TELEFAX: (703)683-4109
7
8 INFORMATION FOR SEQ ID NO: 14:
9
10 SEQUENCE CHARACTERISTICS:
11
12 LENGTH: 7218 base pairs
13 TYPE: nucleic acid
14 STRANDEDNESS: single
15 TOPOLOGY: linear
16
17 IMMEDIATE SOURCE:
18 CLONE: pT5gpt-Fls
19
20 US-08-232-463-14
21
22 Query Match 1.98; Score 47.2; DB 1; Length 7218;
23 Best Local Similarity 5.6%; Pred. No. 0.18;
24 Matches 22; Conservative 206; Mismatches 164; Indels 0; Gaps 0;
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
9
```



```

1 RESULT      8
2 US-09-007-005-17
3 ; Sequence 17, Application US/09007005B
4 ; Patent No. 6258558
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Szostak, Jack W.
7 ; APPLICANT: Roberts, Richard W.
8 ; APPLICANT: Liu, Rihie
9 ; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
10 ; TITLE OF INVENTION: FUSIONS
11 ; FILE REFERENCE: 00786/350003
12 ; CURRENT APPLICATION NUMBER: US/09/007,005B
13 ; CURRENT FILING DATE: 1998-01-14
14 ; EARLIER APPLICATION NUMBER: 60/035,963
15 ; EARLIER FILING DATE: 1997-01-27

```

STREET: 1100 NO. 6140085th Giebe Road, 8th Floor
City: Arlington
STATE: Virginia
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,273
FILING DATE: 01-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCY/GB96/01332
FILING DATE: 03-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9511196.9
FILING DATE: 02-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MS Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-29
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SRD ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 9048 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Brassica
US-08-973-273-4

Query Match 1.98; Score 46.8; DB 3; Length 9048;
Best Local Similarity 52.08; Pred. No. 0.23;
Matches 105; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 124 gctataagtgatgattgaacaccttcttcattgaattagttggaattgacatgata 183
DB 505 GATTATATATATATTTCTAGCTGATGTCGCTAATGCTTATTAACAAATACCTTAT 446
QY 184 gcatatcatatgacatgacgaacattccctacacacacagagaabaatgagagatgta 243
DB 445 GAATTTATATATATTAACGATTTTATTAACAGGAAATATATATGTTATATTTAAT 386
QY 244 atgtaggttgctgttaataataacaaacacacacattagtttttagattttat 303
DB 385 ATAAATATATGATCTCTATATATATATATATATATATCAAAAATTTAAATAGTTAA 326
QY 304 ttatttttataatggtgcta 325
DB 325 TTCTATATATATTTGTTGCTA 304

RESULT 11

US-08-998-416-535
Sequence 535, Application US/08998416
Patent No. 6239264

GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PP/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8689
TELEFAX: 919-541-8587
INFORMATION FOR SEQ ID NO: 535:
SEQUENCE CHARACTERISTICS:
LENGTH: 827 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1373RP
US-08-998-416-535

Query Match 1.88; Score 44; DB 4; Length 827;
Best Local Similarity 46.48; Pred. No. 0.51;
Matches 143; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 1012 tatttaataatattacatgctctatgcttaactatgtaaatggtgaccacttaagta 1071
DB 342 TAATATATAGAT 401
QY 1072 ttatatacatgatacatatcttaaggtacatacaaaatgctacatgaaactttgcaa 1131
DB 402 TATATATATTAATGTAACCTTCAATATATATATATATATATATATATATATAT 461
QY 1132 ttcaatctactgtctcattgtagatgtagctttccacatgctttgaaatgctcgg 1191
DB 462 TAATAGTCTACCTTTAATGATATTAATGATTAATTAATTAATTAATTAATTAAT 521
QY 1192 atctgaattcttaattagatgattgttgggtgcaacgcttaattcttgattatga 1251
DB 522 ATTAAGATATCTTAAT 581
QY 1252 tgtcaaaatcagagcgcttaagaactcttaactaatcttaaaaaatcgattaa 1311
DB 582 TAATATATATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 641
QY 1312 agaaaata 1319
DB 642 TGATATA 649

RESULT 12

US-08-998-416-224
Sequence 224, Application US/08998416
Patent No. 6239264

GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:


```

      1
      2
      3
      4
      5
      6
      7
      8
      9
     10
     11
     12
     13
     14
     15
     16
     17
     18
     19
     20
     21
     22
     23
     24
     25
     26
     27
     28
     29
     30
     31
     32
     33
     34
     35
     36
     37
     38
     39
     40
     41
     42
     43
     44
     45
     46
     47
     48
     49
     50
     51
     52
     53
     54
     55
     56
     57
     58
     59
     60
     61
     62
     63
     64
     65
     66
     67
     68
     69
     70
     71
     72
     73
     74
     75
     76
     77
     78
     79
     80
     81
     82
     83
     84
     85
     86
     87
     88
     89
     90
     91
     92
     93
     94
     95
     96
     97
     98
     99
    100
    101
    102
    103
    104
    105
    106
    107
    108
    109
    110
    111
    112
    113
    114
    115
    116
    117
    118
    119
    120
    121
    122
    123
    124
    125
    126
    127
    128
    129
    130
    131
    132
    133
    134
    135
    136
    137
    138
    139
    140
    141
    142
    143
    144
    145
    146
    147
    148
    149
    150
    151
    152
    153
    154
    155
    156
    157
    158
    159
    160
    161
    162
    163
    164
    165
    166
    167
    168
    169
    170
    171
    172
    173
    174
    175
    176
    177
    178
    179
    180
    181
    182
    183
    184
    185
    186
    187
    188
    189
    190
    191
    192
    193
    194
    195
    196
    197
    198
    199
    200
    201
    202
    203
    204
    205
    206
    207
    208
    209
    210
    211
    212
    213
    214
    215
    216
    217
    218
    219
    220
    221
    222
    223
    224
    225
    226
    227
    228
    229
    230
    231
    232
    233
    234
    235
    236
    237
    238
    239
    240
    241
    242
    243
    244
    245
    246
    247
    248
    249
    250
    251
    252
    253
    254
    255
    256
    257
    258
    259
    260
    261
    262
    263
    264
    265
    266
    267
    268
    269
    270
    271
    272
    273
    274
    275
    276
    277
    278
    279
    280
    281
    282
    283
    284
    285
    286
    287
    288
    289
    290
    291
    292
    293
    294
    295
    296
    297
    298
    299
    300
    301
    302
    303
    304
    305
    306
    307
    308
    309
    310
    311
    312
    313
    314
    315
    316
    317
    318
    319
    320
    321
    322
    323
    324
    325
    326
    327
    328
    329
    330
    331
    332
    333
    334
    335
    336
    337
    338
    339
    340
    341
    342
    343
    344
    345
    346
    347
    348
    349
    350
    351
    352
    353
    354
    355
    356
    357
    358
    359
    360
    361
    362
    363
    364
    365
    366
    367
    368
    369
    370
    371
    372
    373
    374
    375
    376
    377
    378
    379
    380
    381
    382
    383
    384
    385
    386
    387
    388
    389
    390
    391
    392
    393
    394
    395
    396
    397
    398
    399
    400
    401
    402
    403
    404
    405
    406
    407
    408
    409
    410
    411
    412
    413
    414
    415
    416
    417
    418
    419
    420
    421
    422
    423
    424
    425
    426
    427
    428
    429
    430
    431
    432
    433
    434
    435
    436
    437
    438
    439
    440
    441
    442
    443
    444
    445
    446
    447
    448
    449
    450
    451
    452
    453
    454
    455
    456
    457
    458
    459
    460
    461
    462
    463
    464
    465
    466
    467
    468
    469
    470
    471
    472
    473
    474
    475
    476
    477
    478
    479
    480
    481
    482
    483
    484
    485
    486
    487
    488
    489
    490
    491
    492
    493
    494
    495
    496
    497
    498
    499
    500
    501
    502
    503
    504
    505
    506
    507
    508
    509
    510
    511
    512
    513
    514
    515
    516
    517
    518
    519
    520
    521
    522
    523
    524
    525
    526
    527
    528
    529
    530
    531
    532
    533
    534
    535
    536
    537
    538
    539
    540
    541
    542
    543
    544
    545
    546
    547
    548
    549
    550
    551
    552
    553
    554
    555
    556
    557
    558
    559
    560
    561
    562
    563
    564
    565
    566
    567
    568
    569
    570
    571
    572
    573
    574
    575
    576
    577
    578
    579
    580
    581
    582
    583
    584
    585
    586
    587
    588
    589
    590
    591
    592
    593
    594
    595
    596
    597
    598
    599
    600
    601
    602
    603
    604
    605
    606
    607
    608
    609
    610
    611
    612
    613
    614
    615
    616
    617
    618
    619
    620
    621
    622
    623
    624
    625
    626
    627
    628
    629
    630
    631
    632
    633
    634
    635
    636
    637
    638
    639
    640
    641
    642
    643
    644
    645
    646
    647
    648
    649
    650
    651
    652
    653
    654
    655
    656
    657
    658
    659
    660
    661
    662
    663
    664
    665
    666
    667
    668
    669
    670
    671
    672
    673
    674
    675
    676
    677
    678
    679
    680
    681
    682
    683
    684
    685
    686
    687
    688
    689
    690
    691
    692
    693
    694
    695
    696
    697
    698
    699
    700
   
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,179
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Melis, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1909
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5655 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: exon
LOCATION: 2787..3347
OTHER INFORMATION: /product- "1st exon of NIM1"
FEATURE:
NAME/KEY: exon
LOCATION: 3427..4162
OTHER INFORMATION: /product- "2nd exon of NIM1"
FEATURE:
NAME/KEY: exon
LOCATION: 4271..4474
OTHER INFORMATION: /product- "3rd exon of NIM1"
FEATURE:
NAME/KEY: exon
LOCATION: 4586..4866
OTHER INFORMATION: /product- "4th exon of NIM1"
FEATURE:
NAME/KEY: CDS
LOCATION: join(2787..3347, 3427..4162, 4271..4474, 4586..4866)
US-08-880-179-2

Query Match 1.7%; Score 43.4; DB 3; Length 5655;
Best Local Similarity 49.3%; Pred. No. 1.1;
Matches 113; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 1769 agaaaaagcgaacccttaagcattacatgatacagacctcttcttccctt 1828
DB 1654 ATATAAGGTAAATAATTTGCTTCGCGTTTCTTTTACTTTGTTCTTAATGAT 1595
QY 1829 tattttatttctcaggaacttttcttacttaagaacctccaactactaact 1888
DB 1594 AAGTTAAATAGATAAGTTGTAATTTAAAGTAATTAACAATACTCTATAAC 1535
QY 1889 acactcccatgtagaataaagaataataataagaatctgtaactttgtaactagaa 1948
DB 1534 TCAATACACATCATATTTAATTTACTAATTTCTTTGAAACAATTTATGAAAT 1475
QY 1949 aatatatttgcctgttaatttctgtaagftaataacacattttcagt 1997
DB 1474 AGTTTCTTTAATTAATTTTAAATGATATTAATTAATTAATTAAT 1426

Search completed: November 22, 2001, 22:01:01
Job time: 29813 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2001, 13:20:33 ; Search time 1858.82 Seconds
(without alignments)
12723.685 Million cell updates/sec

Title: US-09-578-827A-4

Perfect score: 2502
Sequence: 1 agaagcagaagcgtgggttt.....ttcattaaaaaaaaaattg 2502

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1022815 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*

190: gb_est110:*
 191: gb_est111:*
 192: gb_hic:*
 193: em_gss_fun:*
 194: em_gss_hum1:*
 195: em_gss_hum2:*
 196: em_gss_hum3:*
 197: em_gss_hum4:*
 198: em_gss_hum5:*
 199: em_gss_hum6:*
 200: em_gss_hum7:*
 201: em_gss_hum8:*
 202: em_gss_hum9:*
 203: em_gss_inv1:*
 204: em_gss_inv2:*
 205: em_gss_inv3:*
 206: em_gss_inv4:*
 207: em_gss_inv5:*
 208: em_gss_inv6:*
 209: em_gss_inv7:*
 210: em_gss_inv8:*
 211: em_gss_inv9:*
 212: em_gss_inv10:*
 213: em_gss_inv11:*
 214: em_gss_inv12:*
 215: em_gss_inv13:*
 216: em_gss_inv14:*
 217: em_gss_inv15:*
 218: em_gss_inv16:*
 219: em_gss_inv17:*
 220: em_gss_inv18:*
 221: em_gss_inv19:*
 222: em_gss_inv20:*
 223: em_gss_inv21:*
 224: em_gss_inv22:*
 225: em_gss_inv23:*
 226: em_gss_inv24:*
 227: em_gss_inv25:*
 228: em_gss_inv26:*
 229: em_gss_inv27:*
 230: em_gss_inv28:*
 231: em_gss_inv29:*
 232: em_gss_inv30:*
 233: em_gss_inv31:*
 234: em_gss_inv32:*
 235: em_gss_inv33:*
 236: em_gss_inv34:*
 237: em_gss_inv35:*
 238: em_gss_inv36:*
 239: em_gss_inv37:*
 240: em_gss_inv38:*
 241: em_gss_inv39:*
 242: em_gss_inv40:*
 243: em_gss_inv41:*
 244: em_gss_inv42:*
 245: em_gss_inv43:*
 246: em_gss_inv44:*
 247: em_gss_inv45:*
 248: em_gss_inv46:*
 249: em_gss_inv47:*
 250: em_gss_inv48:*
 251: em_gss_inv49:*
 252: em_gss_inv50:*
 253: em_gss_inv51:*
 254: em_gss_inv52:*
 255: em_gss_inv53:*
 256: em_gss_inv54:*
 257: em_gss_inv55:*
 258: em_gss_inv56:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

RESULT      5
CNS00EVL   1101 bp   DNA      GSS      04-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION     AL069706.1  GI:4949849
KEYWORDS    GSS.
SOURCE      fruit fly.
            Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
            Genoscope.
            Direct Submission
            Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            p1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
    source
        1..1101
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
        /clone_lib="RPCI-98"
        /clone="BACR29B23"
        /note="end : T7"
BASE COUNT   419 a      91 c      60 g      299 t      232 others
ORIGIN
Query Match      3.2%; Score 81.2; DB 219; Length 1101;
Best Local Similarity 36.7%; Pred. No. 1.6e-05;
Matches 217; Conservative 92; Mismatches 280; Indels 3; Gaps 2;
QY 1859 ttaatgaacctccaactctaactaactaccatcgtagaataaagaataata 1918
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 490 TWAATWAAAAAAMWMMWMTTWTWMMWMTWTTWMTWMTWMTWMTWMTWMTW 547
QY 1919 taagatcttgtagatcttgtagtaataataatctgctcgttaatttcgaagc 1978
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 548 TAAATTTAAAMWATWMTTAAAMWMTTAAAMWMTATATTAATWATWATWATA 607
QY 1979 taatcaacattcttcagtagaacaataattacgcacaagaagtagatcattcttg 2038
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 608 TAAAAAAATATTTTWTATATAAATTTTAAATTAATTATTAATTAAATWMTW 667
QY 2039 tccaaatccagtagtagtagtagtagtagtaaaacaaacacacattcgtattgcc 2098
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 668 WMTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 727
QY 2099 caaaaaataaagagaagaagaataatgctcaaaagtgctctctctctcctaataatg 2158
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 728 AAAAAAATTAATAAATAATTAATAATTAATAATTAATAATTAATAATTAATA 787
QY 2159 ttctactaaaccatagatctcaaacagctctcaaaagctcaaaagataaacaatgagac 2218

```

```

DB 788 WATANAATTAATAAATAATTAATAATTAATAATAATAATAATAATAATAATA 847
QY 2219 aacaatcgtatgcaaaaaatcctcttcatgctcttcttctatctcagctttaa 2278
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 848 AAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 907
QY 2279 ttaactaaataatcacaatacccaaacattctctacacaccctcattcgtga 2338
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 908 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 967
QY 2339 ttaccacatccaccagagaacacagaataataataataataataataagac 2398
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 968 TTWMTWMTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1027
QY 2399 aacacatgagctg-atgaatatcacacaagaatataatcagat 2449
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1028 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1079
RESULT      6
CNS00ITT   974 bp   DNA      GSS      03-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence TERT end of BAC:
DEFINITION BAC37D06 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL075432
VERSION     AL075432.1  GI:4954990
KEYWORDS    GSS.
SOURCE      fruit fly.
            Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 974)
            Genoscope.
            Direct Submission
            Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            p1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
    source
        1..974
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
        /clone_lib="RPCI-98"
        /clone="BAC37D06"
        /note="end : TERT3"
BASE COUNT   77 a      59 c      13 g      605 t      220 others
ORIGIN
Query Match      3.1%; Score 77.6; DB 219; Length 974;
Best Local Similarity 33.9%; Pred. No. 7.6e-05;
Matches 285; Conservative 101; Mismatches 454; Indels 0; Gaps 0;
QY 1003 ttctctctatttataatcattacatgctctcattgtaataatgagac 1062
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 TTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 124

```


ACCESSION	AL175696
VERSION	AL175696.1 GI:7813753
KEYWORDS	GSS: genome survey sequence.
SOURCE	Tetradon nigroviridis.
ORGANISM	Tetradon nigroviridis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes; Tetraodontidae; Tetradon. 1 (bases 1 to 1092)
REFERENCE	Roeest-Crolius,H., Jallion,O., Dasilva,C., Fizesen,C., Fisher,C., Bonneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
TITLE	Unpublished
JOURNAL	2 (bases 1 to 1092)
REFERENCE	Roeest-Crolius,H., Jallion,O., Dasilva,C., Bonneau,L., Fisher,C., Bernot,A., Fizesen,C., Wincker,P., Brottier,P., Queller,F., Saurin,W. and Weissenbach,J. Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
COMMENT	Unpublished 3 (bases 1 to 1092) Genoscope. Direct Submission Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-and sequencing project of the tetradon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetradon .
FEATURES	Location/Qualifiers 1..1092 /organism="Tetradon nigroviridis" /db_xref="taxon:39883" /clone="222L11" /clone_1lb="G" /note="Genoscope sequence ID : CGAGC222CF06LP1-end : T7"
BASE COUNT	383 a 169 c 165 g 262 t 113 others
ORIGIN	
Query Match	3.1%; Score 76.6; DB 220; Length 1092;
Best Local Similarity	40.5%; Pred. 0.00012;
Matches 139; Conservative 49; Mismatches 135; Indels 0; Gaps 0.	
Oy 1810	tctcatccttcttcctctattatcttctcagaccttcttctaatagaacc 1869
Dd 997	TTTTTTTAAATAATTAATTAAWTTTNNMTTAATTTTTTTTTTTWTAAAAAA 938
Oy 1870	tccaaactaactaactaacaccccatcgtagaataaagaatatataagaattgt 1929
Dd 937	WWAAAAAAAAAAWATTTAAWMTTWMAAAAAMWMAWMTTWTATTTTATTTT 878
Oy 1930	tgaatttgtagtaacagaataatattgcctcgtaatttcogtaagttaaacaact 1989
Dd 877	TTTAAWTTTAAWTTTAAWMTTWTAAWMTTAAWMTTAAWMTTAAWMTTAAWMTN 818
Oy 1990	tttccagtaagaacaataatctcgcgaagaaglaagatcaatttttgtccaacttc 2049
Dd 817	TTTTTTTAAATAWMAWMTWAATWMTTAAAWMTAAWMAWMTTWTWTWMMWMTWAAA 758
Oy 2050	agttgactatagggtgtgtcgtgtaaaaacaaacacacatcttgattggccccaaaaataa 2109
Dd 757	AAWMAWMTWMTWTWTWTWTWTWTWTWTWMTTAAWMTTWTWTWTWTWTWTWTWTWT 698
Oy 2110	gagagagaagaatattgtccaagaagtgctctctctccta 2152
Dd 697	AAAAAAAAAAATTTTTTTTTTAAAMMMWMTWTWTWTWTWT 655

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2001, 13:28:18 ; Search time 3308.22 Seconds

(Without alignments)
11698.227 Million cell updates/sec

Title: US-09-578-827A-4

Perfect score: 2502

Sequence: 1 agaaagcagcgcgtgsggtt.....ttcattaaaaaaaaaatg 2502

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_cm:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pl1:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_inv3:*
- 33: em_htg_inv4:*
- 34: em_htg_inv5:*
- 35: em_hum1:*
- 36: em_hum2:*
- 37: em_hum3:*
- 38: em_hum4:*
- 39: em_hum5:*
- 40: em_hum6:*
- 41: em_hum7:*
- 42: em_cm:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_stg:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_v1:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_v11:*
- 59: gb_v12:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rol:*
- 95: gb_rol2:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2502	100.0	91740	13 ATF19F18
2	2502	100.0	196339	13 ATCHRIV87
3	1201	48.0	2825	13 AF233752
4	96.2	3.8	7218	10 I66494
5	80.6	3.2	4601	6 DMU11584
6	80.6	3.2	19517	6 DMU37541
7	77	3.1	176174	86 AC007483
8	76.8	3.1	165097	84 CNS057CG

9	76.4	3.1	183250	62	AC012492	AC012492 Homo sapi
10	75	3.0	879	53	CNS01JRG	AL147405 Anopheles
11	72.8	2.9	176174	86	AC007483	AC007483 Homo sapi
12	71.8	2.9	5665	13	AF203688	AF203688 Canavalia
13	71.4	2.9	192929	60	AC005505	AC005505 Plasmid
14	71.4	2.9	185800	81	AL512783	AL512783 Homo sapi
15	71.2	2.8	2305	96	S96842	S96842 SP96-spore-
16	71.2	2.8	7347	2	AF211124	AF211124 Carsonella
17	70.2	2.8	148061	87	AC010727	AC010727 Homo sapi
18	69.2	2.8	94384	87	AC011718	AC011718 Homo sapi
19	69	2.8	183358	87	AC019041	AC019041 Homo sapi
20	68.8	2.7	68709	6	CER683B	AL032855 Caenorhab
21	68.8	2.7	110000	83	CE137H9_0	AL022597 Caenorhab
22	68.8	2.7	253516	83	CEYB3	AL022597 Caenorhab
23	68.8	2.7	330612	83	CEYB7G2	AL022597 Caenorhab
24	68.4	2.7	160401	87	AC009277	AC009277 Homo sapi
25	68.2	2.7	253305	96	PFMAL3B7	AL033559 Plasmid
26	67.8	2.7	184535	86	CNS05ETC	AL355100 Homo sapi
27	67.6	2.7	208546	76	AC083827	AC083827 Homo sapi
28	67.2	2.7	1141	10	AX083744	AX083744 Sequence
29	67.2	2.7	231912	77	AC087566	AC087566 Mus muscu
30	67	2.7	19517	6	DMU37541	U37541 Drosophila
31	67	2.7	165097	84	CNS05ETC	AL355097 Homo sapi
32	66.6	2.7	964	54	CNS07EBR	AL441457 T7 end of
33	66.6	2.7	253305	96	PFMAL3B7	AL034559 Plasmid
34	66.4	2.7	1141	10	AX083744	AX083744 Sequence
35	66.4	2.7	186666	76	AC009409	AC009409 Homo sapi
36	66.2	2.6	156060	60	AC004153	AC004153 Plasmid
37	66.2	2.6	174384	61	AC009524	AC009524 Homo sapi
38	65.8	2.6	4601	81	DMU11584	U11584 Drosophila
39	65.8	2.6	136098	86	AC006970	AC006970 Homo sapi
40	65.6	2.6	6372	6	DDICMPA	M23449 Dictyosellu
41	65.2	2.6	83440	68	AC024285	AC024285 Homo sapi
42	65.2	2.6	185603	74	AC073136	AC073136 Homo sapi
43	65	2.6	170157	79	AL355372	AL355372 Homo sapi
44	65	2.6	174565	74	AC069525	AC069525 Homo sapi
45	65	2.6	178087	85	AC005089	AC005089 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AT19F18 91740 bp DNA
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone F19F18 (EBSA project)
ACCESSION AF035605
VERSION AF035605.1 GI:4468976
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 91740)
AUTHORS Bevan, M., Koester, P., Hempel, S., Entian, K.-D., Bancroft, I.,
Mewes, H.W., Mayer, K.F.X. and Schueller, C.

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 91740)
AUTHORS EU Arabidopsis sequencing project.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mpibiochem.mpg.de, mayer@mpibiochem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk

COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosome 4 can be viewed at: <http://webserver.mips.biochem.mpg.de/Proj/thal/>.

FEATURES
source I. 91740

/organism="Arabidopsis thaliana"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
2053. 3409
/gene="F19F18.10"
join(2053. 2268, 2351. 2551, 2625. 2790, 3003. 3409)
/gene="F19F18.10"
/EC_number="1.11.1.7"
/note="Contains peroxidases signatures [DM1ALSANTL]
contains EST gb:H76577, 235618, H37596, R90036, 235219,
AA713329, AA721954"
/codon_start=1
/product="peroxidase, prx2"
/protein_id="CA838291.1"
/db_xref="GI:446897"
/translation="MVMVNTNLLILLISLITLDSLAQLRRNFYAGCPVEQIV
NAVQKVOQFTTIPATLRIFHDCFNCGDSVMTASTNKKAEKDEHNSLSDG
FDIVAKAEALDVPNCRNKSCADILTMATDVNVLACGPOYDELGRDLSSSTA
SVGKLPHPDDVNTKLSLPAKNGLSLNDMIALSGAHTLGFAHCTVNRITFTT
KDPYTNKDVTELEKASCPENIDPRAINMDPTTPQFNVYKINLQDGKGLFTSDV
LFTDRSKPTVDLMANNGLFNQAFINSMTKIGRVGVTGSGNINIRDCGARN"
2053. 2268
/gene="F19F18.10"
/number=1
2269. 2350
/gene="F19F18.10"
/number=1
2351. 2551
/gene="F19F18.10"
/number=2
2552. 2624
/gene="F19F18.10"
/number=2
2625. 2790
/gene="F19F18.10"
/number=3
2791. 3002
/gene="F19F18.10"
/number=3
3003. 3409
/gene="F19F18.10"
/number=4
5135. 6431
/gene="F19F18.20"
join(5135. 5350, 5437. 5637, 5727. 5892, 6025. 6431)
/gene="F19F18.20"
/note="strong similarity to peroxidase, Arabidopsis
thaliana, gb:X98314
contains peroxidases signatures [DM1ALSANTL]"
/codon_start=1
/product="peroxidase-like protein"
/protein_id="CA838292.1"
/db_xref="GI:4468978"
/translation="MVMVNTNLLILLISLFIATNLSNQLRGDFYAGCPVEIV
FQVIAKAEALDVPNCRNKSCADILTMATDVNVLACGPOYAVELRRDGLSSAS
SVTGKLPKPTFDNLNALFAENGSLSPNDMIALSGAHTLGFAHCTVNRITFTT
NVDPTINKDVTELEKASCPENIDPRAINMDPTTPQFNVYKINLQDGKGLFTSDV
LFTDRSKPTVDLMANNGLFNQAFINSMTKIGRVGVTGSGNINIRDCGARN"
5135. 5350
/gene="F19F18.20"
/number=1
5351. 5436
/gene="F19F18.20"
/number=1
5437. 5637
/gene="F19F18.20"
/number=2
5638. 5726
/gene="F19F18.20"
/number=2
5727. 5892

exon
intron
exon
intron
exon

/gene="F19F18.20"
 /number=3
 5893. .6024
 /gene="F19F18.20"
 /number=3
 6025. .6431
 /gene="F19F18.20"
 /number=4
 complement(join(9998. .10536,10675. .10858))
 /gene="F19F18.30"
 complement(9998. .10536)
 /gene="F19F18.30"
 /number=1
 9998. .10858
 /gene="F19F18.30"
 complement(join(9998. .10536,10675. .10858))
 /gene="F19F18.30"
 /note="similarity to predicted protein, Arabidopsis
 thaliana EST gb:R65200"
 contains EST gb:R65200"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB38293.1"
 /db_xref="GI:4468979"
 /translation="MSCNGCRVLEKSGSENCILPCLQWIESAESQGHATVFAKFG
 RAGLASFISVPELQRPALFQSLFEKGRVNPNGAVGLMTRNMHVCQAAVEYVL
 RGTLRPISDLSEPSLMISDESEIMHDSVSNQTHHCFSTSRSTTEKKSIVNR
 KRLKSDSLDLQVHGLTLTAPAVPPELPPSPFCVKVKGDPGSPSESVTSWEN
 GMRDNNKRNKGEKLLNLFV"
 complement(10537. .10674)
 /number=1
 complement(10675. .10858)
 /gene="F19F18.30"
 /number=2
 14033. .14251
 /gene="F19F18.40"
 /number=1
 14033. .16078
 /gene="F19F18.40"
 join(14033. .14251,14346. .14583,14704. .14834,14990. .15225,
 15334. .15577,15691. .15798,15956. .16078)
 /gene="F19F18.40"
 /EC_number="3.5.1.49"
 /note="strong similarity to formamidase, Methylobolus
 methylotrophus, PIR2:S74213
 contains EST gb:R46402, AA404801, T25222"
 /codon_start=1
 /product="formamidase-like protein"
 /protein_id="CAB38294.1"
 /db_xref="GI:4468980"
 /translation="MAPLSPRVVVKVDLKKRPMOQDPLNRMHPEIPVAVKAGEF
 FVEMIDAMGVINDSDASDINKLVLTTHLSPIVVDVEGVAAGLAVEIC
 NLGPRGDMGPFSGSPRENGCGFTDHPGATATAYWFERITAYSPDIPVPRGTL
 HPGYTGAPSNELRLINDREROLESESVESLTCEIEEGTPENRIENAKRTIPGR
 ENGSGCDIKNLSRGSKITLPVVEGANLSTDMHFSQDGGISFCGALENGSGFLKLC
 EIRNGMOEYLPMPGPTLHNPPIFEIPVPEPSEMLVFEGISVDESGKHDTADATY
 AYKRAVLAIDYLEFEGYSKEOVYLLCCCEGRLSGIVSPNAVATLAIPTAIFDQ
 DIRPKNRVPVGRVRAKPDVLKSTYDGLPLTKNPSSSS"
 14252. .14345
 /gene="F19F18.40"
 /number=1
 14346. .14583
 /gene="F19F18.40"
 /number=2
 14584. .14703
 /gene="F19F18.40"
 /number=2
 14704. .14834
 /gene="F19F18.40"
 /number=3
 14835. .14989
 /gene="F19F18.40"
 /number=3

exon 14990. .15225
 /gene="F19F18.40"
 /number=4
 15226. .15333
 /gene="F19F18.40"
 /number=4
 15334. .15577
 /gene="F19F18.40"
 /number=5
 15578. .15690
 /gene="F19F18.40"
 /number=5
 15691. .15798
 /gene="F19F18.40"
 /number=6
 15799. .15955
 /gene="F19F18.40"
 /number=6
 15956. .16078
 /gene="F19F18.40"
 /number=7
 16078. .17056
 /gene="F19F18.50"
 /number=1
 16838. .18806
 gene
 Query Match 100.0%; Score 2502; DB 13; Length 91740;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 agaagcagaagcgttgggtttcttctaataatgtagaagaacatgacatgagaaattt 60
 DB 59721 AGAAGCAGAGCGCGGCTTCTTCTTAATAATTGTAGAAAGAACTGATCATGAACTTT 59780
 QY 61 gatctaccagaagatgtatgatacctataagaatgtaaatatctactgcatatgtctagcc 120
 DB 59781 GATCTACCAAGAGATGCTATACATCATAGATGTAAATATCTACTGATTTGTCTAGCC 59840
 QY 121 taggctataatgtatgatttgatcattcttcttaattagtttggaattttagcatgat 180
 DB 59841 TAGGCTATAATGTAGATTTGATTCATCTTCTCATTAATTAGTTGGAATTTAGCATGAT 59900
 QY 181 atagatatatctaaatgttccggaattctccatcacatcagaagaatagagattat 240
 DB 59901 ATAGCATATATCTAAATGTGCGGAACCTTCTCATACATACAGAAATATGACAGATTAT 59960
 QY 241 gtaatgtatgttgccttcttaatacagaataacatcatcatattagtttagatttt 300
 DB 59961 GTAATGTATGTGCTTCTTATATATACAAATATACATCATATTAGTTTAGATTTTT 60020
 QY 301 tatttattttttaaagtgctacgtacgttgcgtacaaattatccaaattttagaac 360
 DB 60021 TATTTATTTTAAATAGTGCTACGTACGCGGAGCAATATATTCATTTTGAAC 60080
 QY 361 ttccggattttaacagaaattaaacaatggtgcataagctcggggggtatagaagaattaa 420
 DB 60081 TTCGGGATTTTAAAGGAATTAACAATAGGCGATGACCTCGGGGAGTAAAGATTA 60140
 QY 421 tgccttgcatacgaacaaagagaataatcatgatagctctgacctaagtcgctgtgt 480
 DB 60141 TGCCTTGATCGAGCAAAACGAGAAATCATGATGAGCTTATGCTTAACTGCGTGTGT 60200
 QY 481 taattagaggttcgcatatatacacaacagtagacatatagataataatgacaacacaca 540
 DB 60201 TAATTAGAGGTTCGATATATACATAACAGTAGATATGGAATAATATGAACACACACA 60260
 QY 541 ccaaaaaagtggaataatcctaagtgtagagaataaagtcctccagtggtggagattc 600
 DB 60261 CCAAAAAAGTGGGAATCTAATAATAGTGTAGAGAAATATTAAGTCTCCAGTGGGAGATTCT 60320
 QY 601 aagaagagacatatgaaggtatataagactcctaacaataatggataacttagtgagaga 660
 DB 60321 AAGAAGAGACATATGAGGTATATAGACTCTTAACAAATAATGAGATGATCTTAGAGAGA 60380

OY	661	gggttttaattgaaacaagatgagttttagagaaacaagaaacaagaagcatgcccctaga	720
Db	60381	GGGTTTAAATTGAAACACAGTGGATGTAAGAACAGAAAACAAAGAACACAGCCCTAGA	604400
OY	721	tttcctgagataaataatcacacattgcgtgtltaataaaggtlaagagataatgcacattgg	780
Db	60441	TTTCTGAGATTAAATTACACATCTGCTGTTTATATAGSTAAGAGAAATATGACACATTGG	605000
OY	781	ttggtttcttcaggttaaatctgaaagaaataataatagtaataatttggaaaacttaaa	840
Db	60501	TTGGTTCTTCAGGGTAAATGGAAGAAAAAAAATAGTAAATTTTGAAAAATCTTAAA	605600
OY	841	taagtaaaagagtcataatcagaaagaaagagaaagaaagaaataagtcgacagaaatgg	900
Db	60561	TAGTAAAGAGGATATATATGAGAACAAAGACGAAAAAGGAAAAAATAGTGGCAGACAAATGG	606200
OY	901	agagaggttaggaaggcaaggaacaaatctggagcctttagatgaltgtgaatgcacgcctcag	960
Db	60621	AGAGAGGTTAGAGAGCCAAAGGCAAAATGTGGAGACTTTGATGATGTGATGACAGCCGCTAG	606800
OY	961	ctttcttcacagccctgcgcccaactcaccaacacccatgaacacttctctccatttata	1020
Db	60681	CTTTCTTCACGCCCTGCCTCCACTACCTACACCTATGACCTTCTCTCTATTTTATA	607400
OY	1021	attacattcaaatgctcctatgttactatgtlaaatggtgacacacttaagattatataat	1080
Db	60741	ATTATATTCACATGCTCTATATTACTATGTGTAATGTGTAATGTGACCACTTAAGTATTATAT	608000
OY	1081	catgatatatctcttaagttctatcaacaataatggtcagatgaacacttttgcattcaactc	1140
Db	60801	CATGATATATCTTATATAGGTATCATACACAAATATGTCATGAATCTTTGCCAATTTCAATCT	608600
OY	1141	actgttcattctgtagatgctcagcttctcacatgcttttgaataatagatcgtcgaaat	1200
Db	60861	ACTGTTCATTGTAGATGCTGATGCTTTCACATGTTTGAAATATNGTGTGATCTGAAT	609200
OY	1201	tctttaatagcatgttttctgtgtgttgcacagtttaattcttgcattatgtatgltcaaaaa	1260
Db	60921	TCTTTAAATATACCATGTTTGTGTGTGCACAGTTTAAATCTTGATTAATGATGATGCTCAAAA	609800
OY	1261	ttcagaagcttcagaaaccttaccacttaattctcttaaaataatcgataaagagaaatag	1320
Db	60981	TTCCAGAGGTTCCAGAACCTTACACTATTTCTTAAAAATATTCATTTAAGAGAAAATAG	610400
OY	1321	agcttcaatgcacaagctgtgatagtaaacagtaagtcgcggaatgltctaaacgattatgag	1380
Db	61041	AGTTTTCATGCAACCGTGTATATGATAGTACGATGTGCGGAATGTCTAAACGATTATGAG	611000
OY	1381	tttagtcttttgatggtctagaatattggaattatgtaagtagacatttctaacttttgttaagc	1440
Db	61101	TTTGGTGTGTTTGATTTGGTTAGATATGGATATAGTAGACATTTCACTTTTGTGTTATCTC	611600
OY	1441	tgttgatcttaagatgctgaagagcttcttattatcttaacaacaagtgaagacttggatcga	1500
Db	61161	TGTTGATTTAGAGATGCGTAAGAGAGCTTTTATATTATTACACAGTTGAGACATTGGGATCGA	612200
OY	1501	tagtaacttgaacaaccttggctgtgttctcatgtcaattggccctatatataacaacaactgta	1560
Db	61221	TAGTACTTGAAGAACACTTGGTTGGTTCTCATGTATTTGGCTCATATATTAACAAACATCTGTA	612800
OY	1561	attatatacgaattcttttcggaatttaacgcacatctgtlaagagatatataaacgtgat	1620
Db	61281	ATTATATACGGAATTTTTTCGGAATTTTACCCCAATCTGTATAGATATATATACCTGTCAT	613400
OY	1621	gtcgttttcaaatcatalgtatgaacgaatcacaagtaagtgtcactactcctcaaatatg	1680
Db	61341	GTCGTTTTCAAATTCATATGTATGAACGATCCACGTAATGCTACATCTTCAATATATTG	614000
OY	1681	catgagagagatatgtattcttaaatcttattttgaagaagaataaagaagggaaggtctaac	1740
Db	61401	CATGAGAGAGTATGTATTTTATTAATTTTATTTTGAAGAAAGAAATTAAGAGGAAGGTTTAC	614600

QY	1741	ttgggttgacgcgtctgagaaaaacaaagaagaagcgaaacccactaagccattacg	1800
Db	61461	TTGGGTGATGCATGTGCAAAACAAAAGAAAAGCAAAACCCACTATACCCATTACAG	61520
QY	1801	atacgcacctcttactctcttctctcttacttattttctcagagactttcttcaact	1860
Db	61521	ATATCGACCTTCTTACTTCTTTTCCCTCTTATTTATTTTCTCAGSACTTTTCTACTT	61580
QY	1661	aatgaaaccccccnaactactactacaatacacctcccactgtagaataaagaatatata	1920
Db	61581	AATGAAAACCCCAACTACTACTCACTAACAATACCTCCCACTGATGAATTAAGAAATTTATA	61640
QY	1921	agatattgttgatattcttgtaactagaaaatataattgcctctgtaatttttcgtaagta	1980
Db	61641	AGATATTGTTGATTTTGTATTTTGTACTAGAAAATATATTTTGCCTGTGAATTTTGTAAAGTTA	61700
QY	1381	aatcaacatttttcagtagaanaacaaatattactgcgaaaagtagaalcattattctgc	2040
Db	61701	AATCAACATTTTTCATAGTAAGAAACAAATATATCTGCAAAAAGTAGATATTTTTGTGC	61760
QY	2041	caaaatctcagtagtataggttgatagggtgtgtgtataaaaacaaacaactctgattcgcca	2100
Db	61761	CAAAATCTCAGTAGTATAGGTATAGGGTTGTGTGTAAAAACAAACACATCTTATTTGCCCA	61820
QY	2101	aaaaataaaagaagaagaatattgttcacaaagtgtctctctctctctcttaatttgt	2160
Db	61821	AAAAATTAAGAGAGAGAAATAATTTGTTCAAAAGTGTCTCTTCTCTCTTAATTTATGTT	61880
QY	2161	tctactaaacccaattagattcaaacagctctacaagctccaaagaataacaactcagacaa	2220
Db	61881	TTCACTTAACCCCATTTAGATTCMAACAGTTCACAAAGTCMAAAGATCAAAAGATTAACATGGACAA	61940
QY	2221	caatttgatgcaaaaaaattcctcttcttactgtctcttttttactctctagctcttaatt	2280
Db	61941	CAATTGATGTCAAAAAATCTCTTTTCAATGCTCTTTTATTTCTATCTTAGCTTPTTAAAT	62000
QY	2281	actaatataaaactcacaacatccacaaacccattctctacaaactcaccactctactagatt	2340
Db	62001	ACTAATATAAAACTCACAAATTCACCAAAACCATTTCTTACAACTCACCTTCATCTTAGATT	62060
QY	2341	taccacactccacacgagaacaacagaanaaaaaatatataataataataaagaacaa	2400
Db	62061	TACCCACTCCCAACCGAGAAACACAAAGAAAAAATATATATATATAATATACAGSACA	62120
QY	2401	caaatgattgctgtagtaatatatacacaaacaaagtcttaaatcttagatatgtggtctcc	2460
Db	62121	CACATGATGCTGATGCAATATTCACACAAAGATTATAAATTATAGATATTGTGGGTCTCC	62180
QY	2461	ctctctctatcattcttcttcttacttataaaaaaanaaaatg 2502	
Db	62181	CTTCTCTCATTCATTTTCTTATTTCTATTAATAAAAAAATAAATG 62222	

	RESULT	2
LOCUS	ATCHR187	
DEFINITION	ATCHR187 196339 bp DNA	PLN
ACCESSION	AL161591	16-MAR-2000
VERSION	AL161591.2	contig fragment No. 87.
KEYWORDS		
SOURCE	thale cress.	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis	
AUTHORS	1. (bases 1 to 196339) Rose, M., Hempel, S., Entlan, K.-D., Nemes, H.W., Lemcke, K. and Mayer, K.F.X.	
JOURNAL	Unpublished	
REFERENCE	2. (bases 1 to 196339)	
AUTHORS	EU Arabidopsis sequencing/project.	
TITLE	Direct Submission	
JOURNAL	Submitted (10-MAR-2000) WTPS, at the Max-Planck-Institut fuer	

COMMENT	
<p>Biochemie, Am Kioferspitz 18a, D-82152 Martinsried, FRG, E-mail: blochem@mps.blochem.mpg.de, mayer@mps.blochem.mpg.de Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.blochem.mpg.de/proj/thal/ this fragment has an overlap with ATCHRIV88 at the 5' end and an overlap with ATCHRIV88 at the 3' end.</p>	
FEATURES	
source	Location/Qualifiers
gene	1..196339
	/organism="Arabidopsis thaliana"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
	6146..7792
	/gene="At4g37210"
	6146..6474
	/gene="At4g37210"
	/number=1
CDS	join(6146..6474,6702..6919,7013..7225,7313..7403,7489..7792)
	/gene="At4g37210"
	/note="intron number 1 is a special U12 intron similarity to nuclear histone-binding protein NL/N2, Xenopus laevis, PIR2:A25680"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAB80387.1"
	/db_xref="GI:7270704"
	/translation="MVEESASASASVITQTEPATRIAQTLPEPNLSTIETVSYVQGGESTNNDANNANADSAATEVCDEDERKTEFAELTEKGSVFLKENDFAAVDCESRLAEIRVAHGEGLDAECINAYRYGLALANKQAADPLGNMPKKEGVQOESSNGESLAPSVSGSDPEROGSSGSGGKDGEDCDLSDADGADDESDLDNAWKMDIARVTDKOSTEMEKVILCSLAESLEREDIESLSQYKALSTIELVPSDREARINRICLETGCGPKRAIPYCKKAILKARMERLSNETIKGASGATSTSVSEIDGIIQSSNPIYIDKSAKDEKVEIGDLAIEDLEKAKSLNLSVH"
	6475..6701
	/gene="At4g37210"
exon	/number=1
	6702..6919
	/gene="At4g37210"
	/number=2
	6920..7012
	/gene="At4g37210"
	/number=2
	7013..7225
	/gene="At4g37210"
	/number=3
Intron	7226..7312
	/gene="At4g37210"
	/number=3
	7313..7403
	/gene="At4g37210"
	/number=4
	7404..7488
	/gene="At4g37210"
	/number=4
	7489..7792
exon	/gene="At4g37210"
	/number=5
	8876..9739
	/gene="At4g37220"
	join(8876..9051,9141..9248,9330..9400,9486..9739)
	/gene="At4g37220"
	/note="strong similarity to cold acclimation protein WCO413, <i>Trifolium aestivum</i> , PATCHK:G165785 Contains Prokaryotic membrane lipoprotein lipid attachment site AA147-157 contains EST gb:AW033651.1, W43270, AA650647, A1966990.1, AA728669, T42795, Z37671, A1100742, T42949, AA040998, AA395771, AA657303, T41871, T45633"
	9051..9140
	/gene="At4g37220"
	/number=1
Intron	9052..9140
	/gene="At4g37220"
	/number=1
	9141..9248
	/gene="At4g37220"
	/number=2
	9249..9329
	/gene="At4g37220"
	/number=2
	9330..9400
exon	/gene="At4g37220"
	/number=3
	9401..9485
	/gene="At4g37220"
	/number=3
	9486..9739
	/gene="At4g37220"
	/number=4
	10194..11357
	/gene="At4g37230"
CDS	10194..10398
	/gene="At4g37230"
	/number=1
	join(10194..10398,11131..11357)
	/gene="At4g37230"
	/note="possible frameshift in DNA sequence at pos. 51188-51200, 5' part of gene couldn't be reconstructed, possible pseudogene, no ATG strong similarity to photosystem II oxygen-evolving complex protein 1, spinach, PIR2:A23626 contains EST gb:z34685"
	/codon_start=1
	/product="photosystem II oxygen-evolving complex like protein (partial)"
	/protein_id="CAB80389.1"
	/db_xref="GI:7270706"
Intron	/translation="MSTLTATLDEIEGPEFVDYAAYVHNFLGVSYSRSGSSSMH RRGTRGKSIPTSAFTGCHGKSDQIGINAYALPARGNNEELAKENKITTLSYTKS NPESGEVIGAFESIQPSDTDLGATTPDKVKIIGIWCQIDE"
	10399..11130
	/gene="At4g37230"
	/number=1
	11131..11357
	/gene="At4g37230"
	/number=2
	18388..18822
	/gene="At4g37240"
	/number=1
exon	18388..18822
	/gene="At4g37240"
	18388..18822
	/gene="At4g37240"
	/note="contains EST gb:z18456"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAB80390.1"
	/db_xref="GI:7270707"
	/translation="MEFANPVKGYLLKYPKFCICNSDDMFDDAAVAISADELQI LGOIFALPLCMLRQPLKAEMALAVASSALMGSGGCGCRKCEPIVSDKILMRVG SGDTVGSGRRKVRNGDGGGVSSSRRRKCYAAELSTIDE"
	21559..23955
	/gene="At4g37250"

exon complement(21559..22253)
/gene="AT4937250"
/number=1
complement(21559..22253,22350..23955)
/gene="AT4937250"
complement(21559..22253,22350..23955))
/gene="AT4937250"
/note="similarity to protein kinase TMK1, receptor type precursor - Arabidopsis thaliana, PIR1.J01674.
contains EST gb:H76836"
/codon_start=1
/product="receptor kinase-like protein"
/protein_id="CA80391.1"
/db_xref="GI:7270708"
/translation="MELISVIFEEFFCSVLSSALNSDGLVLMKFSVLYVPLSLIOT
KMYKHPSCSMRIGISCNDSKIVITLISLPISOLLSIPDLGSLITLQSLDLSNNSNG
PLPVSEFNARELRLDLSNMTSGEIPSLIGLHMLTINLSDNALAGLPINLALR
NITVYSLENNYFSGELPGGRVYEPIDLSNMLNGSLPDEGGISLYQYLVSEFNDISG
ELPELGVNPNVTVDSFNNLTGPIPDSPVFNQESNFSGNPLGCEPRNPLI
PSPSIVSEADVPSTPAIAIPNTIGSNPDVDPNOCQDPMPRTGLRPGVIGIYV
DIAGIGILAVILYLYIRCKKNIYDNNNDKORTEDTITLSTFSSSSSPESRER
KMSCLRKDETPSEEDDEDDDESGYANORSGDNKLVTDGKEMETLILKASA
YLIGATGSSIMKAVLEDGRVRAVRRLGNGCSOREPKRPHIRAIQGLVIRNVLRL
CGFYKCTDEKLYIYDPVNGSLVNPRIKRGSSSPTFLPMTIRLKIAGVIRGLAYL
HEKRVHGLKPSNITLGHDMERKIDFLESLTGETSYIRAGSSRIFFSKRYTTS
SREFFSIGTPSPSSVGAMSPYCAPSEFRSLKSPKMDVYGFVYLLELTGKIVS
VEIYVINGLTVEDGHRAYMADVAIRGELDGKQEFLLDCFKLIGYSCASAPVOKRPTM
KESLAVLERFHNSSVIRKSSFFHYGH"
complement(22254..22349)
/number=1
complement(22350..23955)
/gene="AT4937250"
/number=2
34372..35334
/gene="AT4937260"
34372..35334
CDS

Query Match 100.0%; Score 2502; DB 13; Length 196339;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaaagagagcgggggtttcttctataatctttagaagaactgcatgagaacattc 60
DB 183143 AGAAGCAGACGCTGGGGTCTTCTATATATTTAGAAAGTGTAGACAGAAACATT 183202

QY 61 gactaacagagatggtgatgactacataagatgtaaatatctactgcatatctcagcc 120
DB 183203 GATCTACAGAGATGCTGATGATCATTAAGATTAATCTACTGATATATCTTAGCC 183262

QY 121 taagctataatgtagatcttgatcacttctctcaataatgattggaatttagcatgat 180
DB 183263 TAAGCTATAATGTAAGATTTGATCATCTTCTTCAATTAATGTTGGAATTTAGCATGAT 183322

QY 181 atagcatatacttaaatatgctcgaagaacttctctacactagaaataatgagagattat 240
DB 183323 ATAGCATATATCTTAATATGTCGAAACCTTCTACTACTAGAAATATAGGAGAGATTAT 183382

QY 241 gtaatgtaggttgctgtgtaatacaaaaataacatcatcatgatttagattttt 300
DB 183383 GTAAATGAGGTTCGCTGTAAATACAAAATACATCATCATTTTACGATTTT 183442

QY 301 taattattttataatggtgctacgtagctggtcgaataatattccaattttgagac 360
DB 183443 TAAATTAATTTTAAATAGTGTACGTACGTGCGCATCAAAATTAATCCAAATTTTGAGAC 183502

QY 361 ttgggaatttaaaagaataaagaatgagcgtgagctcggggggagtagaagaattaa 420
DB 183503 TTGGGGAATTTTAAACGAATTAACATAGGCAATGAGCTCGGGGGGATAAAGATTAA 183562

QY 421 tgccttgatcgaagaacagaaataatcatgtagagcctatgcataaagtgccttggt 480
DB 183563 TGCTTTGATACGAGACAAAGAAAATCATGATGAGCCTATGCAATTAAAGTCCGCTTGGT 183622

QY 481 taattgaggttcgcataatacaaaacagtagacatgataatgataatgaaacacaca 540
DB 183623 TAATTAAGAGGTTCGATATATCAATAAACAGTAGACATATGATTAATGACACACACA 183682

QY 541 ccaaaaaagtgggaatctcaataatgctgtagagaataaagtctcaggtgagattc 600
DB 183683 CCAAAAAAGTGGGAATCTTAATTAAGTATGAGAAATAATAGCTCAGAGTGGAGATTTC 183742

QY 601 aagaagaaggaatggaagggtatatagtctcaaaaaaagtgcattagtgagaga 660
DB 183743 AAAGAGAGCAATGAAGGATATATAGCTTAACAAAAATGCGCTAGGTGAGAGA 183802

QY 661 gggctttaaattgaacaagtaagatggaagaacaagaacaagaacatgcccata 720
DB 183803 GGGCTTTAAATTAATTAACAGATAGATTGAAGAACAGAAAAACAGACATGCTCCCTAGA 183862

QY 721 ttctcgaataataatcacatctgctgttataaaggtaagagataatgacacattg 780
DB 183863 TTTCTGAGATAATTAATTAACATATGCTGTTATATAGGTAAGAGAAATGACACATTGG 183922

QY 781 ttggttctcaggggtaaatgtaaaaaaataatgataatcttgagaaatctaaaa 840
DB 183923 TTGGTTCTTACGGGTAAATGTGAAGAAAAAATAATGATTAATGAGAAATCTMAAA 183982

QY 841 tagtaagaagtaataatgtagaagaagaagaagaagaagaagaatgtagcagagattg 900
DB 183983 TATTAAGAAGGTATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 184042

QY 901 agagaggttagagagcaaaagcgaatgtagagcttgatgattgtagcagccgtcag 960
DB 184043 AGAGAGGTTAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 184102

QY 961 ctttctcagcgtctccacactacacacactgaaacattctctcatttata 1020
DB 184103 CTTTCTCAGCGCTCTCCACACTACACACTACACACTGATGAAATCTCTCTATTATTA 184162

QY 1021 attataatcaatgctcctatgtaactgtaactgtaaatgtagcacttaattatata 1080
DB 184163 ATTAATTAATCAATGCTCTATATGATGATGATGATGATGATGATGATGATGATGAT 184222

QY 1081 catgtataatcttaagtagatatacaaaaatgtagcaaaactttgcaattcaatc 1140
DB 184223 CATGTATATATCTTATAGGATCATCAAAAATGTCATGAAACTTTGCAATTTCAATCT 184282

QY 1141 actgttaactgtagatgtagcttcaaatgatttgaataatgtagcagatcgaat 1200
DB 184283 ACTGTTAATGTAAGATGCTAGCTTTTCAATGTTTGAATAATGATCTGATCTGAAAT 184342

QY 1201 tcttaattagcatcttctgtgtgcaagctttaaattcttgattatgtagtcaaaa 1260
DB 184343 TCTTTAATTAAGCATTTGTTTGTGTGTCACACGTTTAATTTCTTGATTAATGATGTCAAAA 184402

QY 1261 ttcaagagctcgaacacttaccataatcttcttaaaaaataatcgatlaagaanaatag 1320
DB 184403 TTCAGAGCGCTTCAGAACTCTTACACTAAATTTCTTAATAATTAATGATTAAGAAAAATAG 184462

QY 1321 agtttcaatgacacagtgtagtaagtagtagtcgagaaagtctcaaaacgattatgag 1380
DB 184463 AGTTTCAATGACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 184522

QY 1381 ctctggtcttgatcttgtagaattgtagaattgtagaagatctcaacttttctgtaagc 1440
DB 184523 TTTGCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 184582

QY 1441 tctgattcaagatcgcgaagaagctcttcttataacacagttgagactggagatcga 1500
DB 184583 TGTGATTATTAAGATGCTGAAGAGTCTTTTATTTATACACAGGTTGAGACTTGGGATGCA 184642

QY 1501 tagtacttaaaacacttggttggttcaatgatttggtcctatatataaaacaacatcgt 1560
DB 184643 TAGTACTTAACACACTTGCTGCTTCAATGATTTGGCCCTATATATAAACAACATCGTA 184702

QY 1561 attataacgattttttcgcgaattttacgcacatctcgttaagtatatataacatgcat 1620

REFERENCE	1 (bases 1 to 7218)
AUTHORS	Dorner, F., Scheifflinger, F. and Falkner, F. Gunter.
TITLE	Recombinant fowlpox virus
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;
FEATURES	Location/Qualifiers
source	1..7218
BASE COUNT	/organism="unknown"
ORIGIN	1944 a 1491 c 1486 g 1929 t 368 others

OY	514	aacatgtgtaaatatgcgaacacacacccaaaagtggaatactcaatatggttagag	573
Db	1477	ACCTATTCTATGCAAGTGTGTTAAAGACTGAGAATAATTGGTCRRRRRRRRRRRRRRR	1418
OY	574	aaataaagtcttcgaatggcgtgagatccaagaagagacaatgaaggtatataactcta	633
Db	1417	RRR	1356
OY	634	aacaaaaatggcatgaccttggtagagagggtttaaatgaaaccaagtggattgaaga	693
Db	1357	RRR	1298
OY	694	caagaacaacaagaagatgccttaagttctctgaataaataaacatgctgttat	753
Db	1297	RRR	1238
OY	754	ataaggtcaagaagaatgatgacacattggtgttctcttacggtlaaatgtaagaaaaa	813
Db	1237	RRR	1178
OY	814	aattgtataatttgagaanaatccaataatgtaaagggatatalatcgagaagaagaaga	873
Db	1177	RRR	1118
OY	874	aaagggaaaaatgctgvcagagaagaatcgagaaggtltagaggccaagycnaatgtgagc	933
Db	1117	RRR	1058
OY	934	ttagatgattgtagtcagc	954
Db	1057	TCCCTGCACCTGCACGCCAACG	1037

RESULT	5				
LOCUS	DMU11584/c				
DEFINITION	DMU11584	4601 bp	DNA	INV	23-JUL-1994
VERSION	011584				
KEYWORDS	U11584.1	GI:508826			
SOURCE	mitochondrial DNA; A+T region; tandem repeats.				
ORGANISM	fruit fly.				
REFERENCE	Mitochondrion Drosophila melanogaster				
AUTHORS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
TITLE	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
JOURNAL	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
MEDLINE	1 (bases 1 to 4601)				
REFERENCE	Lewis, D.L., Farr, C.L., Farguhar, A.L. and Kaguni, L.S.				
AUTHORS	Sequence, Organization and Evolution of the A+T Region of				
TITLE	Drosophila melanogaster Mitochondrial DNA				
JOURNAL	Mol. Biol. Evol. 11, 523-538 (1994)				
MEDLINE	94285822				
REFERENCE	2 (bases 1 to 4601)				
AUTHORS	Kaguni, L.S.				
TITLE	Direct Submission				
JOURNAL	Submitted (28-JUN-1994) Laurie S. Kaguni Ph.D, Dept. of				
MEDLINE	Biochemistry, Michigan State University, East Lansing, MI,				
REFERENCE	48824-1318, USA				

ORGANISM Mitochondrion *Drosophila melanogaster*
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; *Drosophila*.
REFERENCE 1 (bases 12511 to 12682)
AUTHORS Clary,D.O., Goddard,J.M., Martin,S.C., Fauron,C.M. and
Wolstenholme,D.R.
TITLE *Drosophila* mitochondrial DNA: a novel gene order
JOURNAL Nucleic Acids Res. 10 (21), 6619-6637 (1982)
MEDLINE 83090428
REFERENCE 2 (bases 5269 to 5695)
AUTHORS Clary,D.O., Wahlertner,J.A. and Wolstenholme,D.R.
TITLE Transfer RNA genes in *Drosophila* mitochondrial DNA: related 5'
flanking sequences and comparisons to mammalian mitochondrial tRNA
genes
JOURNAL Nucleic Acids Res. 11 (8), 2411-2425 (1983)
MEDLINE 83220794
REFERENCE 3 (bases 404 to 5272)
AUTHORS de Bruijn,M.H.
TITLE *Drosophila melanogaster* mitochondrial DNA, a novel organization and
genetic code
JOURNAL Nature 304 (5923), 234-241 (1983)
MEDLINE 83245048
REFERENCE 4 (bases 804 to 1778)
AUTHORS Satta,Y., Ishiwa,H. and Chigusa,S.I.
TITLE Analysis of nucleotide substitutions of mitochondrial DNAs in
Drosophila melanogaster and its sibling species
JOURNAL Mol. Biol. Evol. 4 (6), 638-650 (1987)
MEDLINE 88174373
REFERENCE 5 (bases 5268 to 13619)
AUTHORS Garesse,R.
TITLE *Drosophila melanogaster* mitochondrial DNA: gene organization and
evolutionary considerations
JOURNAL Genetics 118 (4), 649-663 (1988)
MEDLINE 88212147
REFERENCE 6 (bases 441 to 2967)
AUTHORS Satta,Y. and Takahata,N.
TITLE Evolution of *Drosophila* mitochondrial DNA and the history of the
melanogaster subgroup
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)
MEDLINE 91088537
REFERENCE 7 (bases 14215 to 14512)
AUTHORS Ballard,J.W., Olsen,G.J., Faltth,D.P., Odgers,W.A., Rowell,D.M. and
Atkinson,P.W.
TITLE Evidence from 12S ribosomal RNA sequences that onychophorans are
modified arthropods
JOURNAL Science 258 (5086), 1345-1348 (1992)
MEDLINE 93088057
REFERENCE 8 (bases 14917 to 19517)
AUTHORS Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.
TITLE Sequence, organization, and evolution of the A+T region of
Drosophila melanogaster mitochondrial DNA
JOURNAL Mol. Biol. Evol. 11 (3), 523-538 (1994)
MEDLINE 94285822
REFERENCE 9 (bases 1 to 408; 13319 to 19517)
AUTHORS Lewis,D.L., Farr,C.L. and Kaguni,L.S.
TITLE *Drosophila melanogaster* mitochondrial DNA: completion of the
nucleotide sequence and evolutionary comparisons
JOURNAL Insect Mol. Biol. 4 (4), 263-278 (1995)
MEDLINE 96423163
REFERENCE 10 (bases 1 to 19517)
AUTHORS Lewis,D.L., Farr,C.L. and Kaguni,L.S.
TITLE Direct Substitution
JOURNAL Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department,
Michigan State University, East Lansing, MI 48824-1319, USA
FEATURES
source 1. 19517
/organism="Drosophila melanogaster"
/organism="mitochondrion"
/db_xref="taxon:7227"
/note="derived from new and previously submitted
sequences; sequence is a composite containing sequences
obtained from different *Drosophila melanogaster* strains"

trna 1. 65
/product="trna-1le"
complement(97.165)
trna /product="trna-gln"
171. 239
trna /product="trna-phe"
240. 1265
CDS /codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 2"
/protein_id="AAC4781.1"
/db_xref="GI:1166530"
/translation="MENNSSKILFTITMIIGTLITVTSNMLGAMGLIEIULSPIPL
LSDNNIMSTESTASKYPLIJOVASIVLIPSSILMLKNNNNNEINESTSITNSAIL
LKSQAPFHFEFPNMEGIMKNALMLTQWOTAPLMLISTLNLITLISVILSVIL
GAIQLNOTSEKLMKAFSSINHLGMLSLMSISEIWLIFEEFYSFEVLFEMENIR
KLHFNOLFSMEKNSKILKFTLFNMLSLIGLPPPLGFLPKMLVIOQLTNOYFMTL
IMMSTLITLFFYLIRICYSAPFMNMFENNMIKMNMSININMIMMFFSIFGLFLI
SLFFYFMP"
1264. 1329
trna /product="trna-trp"
complement(1322.1383)
trna /product="trna-cys"
complement(1403.1468)
trna /product="trna-tyr"
join(1470.1472,1474.3009)
CDS /note="mechanism underlying reading frame shift after
first codon uncertain"
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AAC4781.2"
/db_xref="GI:7412849"
/translation="MSRQLFSTNRHIDGTLVFLFGAGAVGTSLSILIRAEIHPG
ALIGDQIYNVIVTAHAFIMIEFTVTPIMIGFGMLVPLMIGADPAFPMNNSEFV
LIPALSLIVSVNENGAGCGWTVYPPISAGIAGASVDLAIFSLIAGISSTIGA
VVFITVIVNMRSTGISLDRMPLFVWSVITALLISLIPVLGATMTLITBNNTSF
PPAGCGDPIIYOHGIPFPGHEVYIILPGGMISHISQESCKEFGSLGMIYAM
LAIGLIGFVNAHHNFTGMVDVTRAITFSATMTAVFGIKITSWLTLTGTOISTIS
PAIHLALGFVLETVGLGLGVLANSSVDIILHDYVVAHFHVLNAGVAFALMAGF
IMVPLFTGLTLNNMKLSHFTIMFGVNLFFPHQFGLGAMPKRYSDYDAYVTNN
IVSTIGSTISLIGLIFFFIILWESIVSOROVYIYQLNSSIEMVQNTPPAHSSVSLP
LITN"
3012. 3077
trna /product="trna-leu"
3083. 3767
CDS /note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except="(pos:3767,aa:TERM)
/transl_table=5
/product="cytochrome c oxidase subunit II"
/protein_id="AAC4781.1"
/db_xref="GI:1166532"
/translation="MSTWANLGLDQASAPLMBOLIFPDHALLITVMTVGYLMEW
LEFNWYVRSFLLIGOLIMYITPAILLIPALSLIVSVNENGAGCGWTVYPPISAGI
OYKWSYVSDRNIEFSDMIPNTEMLTMDGRLLVDNRVYVLPMSQIRIIVTADVI
HSWYPALGVKVDGTPGRINDTNEFINRPGFLFYGCSEICGANNHFMPIVIESVFN
FLKWSNNNS"
3768. 3838
trna /product="trna-lys"
3840. 3906
trna /product="trna-asn"
3907. 4068
CDS /codon_start=1
/transl_table=5
/product="arfpase 8"
/protein_id="AAC4781.1"
/db_xref="GI:1166533"
/translation="MPQMAPISWLLLIITFSITFLFCSINYSYMPNSPKSNELKNI
NINSMNKKW"
4062. 4736
CDS /codon_start=1

/transl_table=5
/product="ATPase 6"
/protein_id="AAC47815.1"
/db_xref="GI:1166534"
/translation="MNTNLSVEDPLAIFNFSNLWSTFLGLMIPSIWIMPSRYNI
MMSNLTLLHKEFKTLPSGHNSTFIISLSLIFNNMGLEPIYESTSLTLT
LSLNPALICFMLGINTOHMFALHVPDGTALIMPFWCIEITINIIIPGLAAR
LTAINGHLLTLTGNTGSSMYLMTFLMQIALLVESANAMQSYFANVLSLT
YSSRYN"
4736..5524
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit III"
/protein_id="AAC47816.1"
/db_xref="GI:1166535"
/translation="MSTHSNHPHLVDYSPMLTGAIGAMTVSGNWKWHDYDISLE
VLGNITILVYQMMRDVSRREGTQGLHTAVTIGLWGMITFLSEVLFVSFPMAR
FHSLSPAIEIGASMPGIIISNPDIPLNTAIIILASGVTTVAHSHLSMENNHSOT
TGLSEFTVLGIVFTIIOAYEIEAPFIDISITGFEPMATGFGHIVLIGTFELLY
CLRLHNNHPSKNHFGFEAAAWYHVDVWVLELYITIKWGS"
5543..5607
/product="tRNA-Gly"
5608..5961
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 3"
/protein_id="AAC47817.1"
/db_xref="GI:1166536"
/translation="MFSIIFALILITITVIMPLASILSKALIDREKSSPECGFD
PKSSRLPEFSRFLITIIIFLIDVEIALIPMIIMKYSNIMWTITSIIIFILILI
GLYHNMGMNMSN"
5981..6045
/product="tRNA-Ala"
6055..6118
/product="tRNA-Arg"
6119..6183
/product="tRNA-Asn"
6184..6251
/product="tRNA-Ser"
6252..6318
/product="tRNA-Glu"
complement(6317..6401)
/product="tRNA-Phe"
complement(6401..8124)
/note="TAA stop codon is completed by the addition of 3' A
residue to the mRNA"
/codon_start=1
/transl_except=(pos:complement(6401..6402),aa:TERM)

Query Match 3.2%; Score 80.6; DB 6; Length 19517;
Best Local Similarity 44.1%; Pred. No. 0.00073;
Matches 566; Conservative 0; Mismatches 704; Indels 13; Gaps 5;

DB 1037 tctatgctactatgtaaatgctgacacctgaagtatttatatacatgtatatacttat 1096
DB 18114 TATATTTTCAATTAATAAATTTATTAATAAATTTTGTATTTTAAATAAACAATGA 18055
OY 1097 aggtatcatcaaaatgctgacaccttgcaattctcaactctgctgctatgtaga 1156
DB 18054 TTTTATATATTAATTTTATTAATAAATATACATTATTAACAATTTTAAATAATTTA 17995
OY 1157 tgcctagcttccacatgtttgaaatagctgcgacatccttaacttaactgcatg 1216
DB 17994 TATTAATTAATTAATTAATTTTCTATATATATATATATATATATATATATATATC 17935
OY 1217 ttgttggtgcacagcttaattcttgattatgtagtgcacaaatctcgagcgctcagaa 1276
DB 17934 AATTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 17875
OY 1277 cctctacacaaatctcctcaaaatcatcgatcaagaagaatagagcttccagccag 1336
DB 17874 AATTAATTAATTAATTAATTAATTAATTAATTTTATTAATAATTAATTAATTAATTT 17815

OY 1337 tgtgatagtaacgtagtcgcggaaatgctcaaacgattatggttgcttgattg 1396
DB 17814 TAAATTTAAATAATTTTATTTAAATTTTATGATATATATATATATATATATATATATAT 17755
OY 1397 gtagaattgtagtagtagacatctcaactcttttgtagctgtagtagtagtagc 1456
DB 17754 TATATATTAAT 17695
OY 1457 gtaaa--gagctcttctatcttcaacagctgtagacttgtagtagtagtagtagc 1514
DB 17694 TTTTATTAATAAATAATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 17635
OY 1515 ctgttggttgcctatgtagtggcctatataaaacacacgtaactatataatcgatt 1574
DB 17634 ATTAATTAATTAATAAATAATTTTGTATTTTAAATAAATAAATAAATAAATAAATAAATA 17575
OY 1575 ttttcggaattctaacgcacatcgtgaagt-----atacatcacgtagcgtttt 1628
DB 17574 TTTTATTAATAAATAAT 17515
OY 1629 caaattcatatgtagaacgtagcagtagtagtactctccatcaatattgtagtagag 1688
DB 17514 TAAATTAATTTTCTAT 17455
OY 1689 agatagtagtattataaatttattttagaagaagaataagaaggagtagtcttggtg 1748
DB 17454 AATAAT 17395
OY 1749 atcgatgtagaaacaaagaagaagaagaaacccactaagccatcacatgtagtagc 1808
DB 17394 ATTAATTAAT 17335
OY 1809 ctcttactcttctccctcttacttatttcttctcaggaacttttctcactaagtagaac 1868
DB 17334 TTATTAATTAAT 17275
OY 1869 ctccaaactatcactatcaactccatccatgtagaataagaataatataata-agaatt 1927
DB 17274 TATTAATTAATAATTTTAAATAATTTTCTATTTTAAATAATATATATATATATATAT 17215
OY 1928 gttagatattgtagaactgaacaaatattgctcgttaatttctcgttagtagtagtagc 1987
DB 17214 TTTTATTAATTTTAAATAATTTTATTTTATATATATATATATATATATATATATAT 17155
OY 1988 attttcagtagaacaacatacttactgcacaaagaagtagatcattatttgcacaaatc 2047
DB 17154 TAAGTATATATTTAATAATTAATTTTATTTTAAATAAATAATTTTAAAGTTTAAATTA 17095
OY 2048 tcaagtagctataggggttgtagtaaaaca--aaacacattcgtgattgcccacaaa 2104
DB 17094 TATATTAATTTATTAAGATAGGCGGAATTAATTTTCAATTTTCAATTTTCAATTTATAT 17035
OY 2105 ataaagaagaagaagaatcatggttcaaaagtgctctcctcctcctcctaatatggtttca 2164
DB 17034 ATATATATATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 16975
OY 2165 ctaaacccaattagatctcaacacagctcaacaaagctcaacaaagtagaagaacacacat 2224
DB 16974 TTAATTAATTTAATTTGAATAATTTAATTTATATGATATATATATATATATATATATAT 16915
OY 2225 tcatgtagcaaaatcctcttctcattcagctcttcttcttctcctcctcctcctcctcctc 2284
DB 16914 T-TTATTAATAAATAATCATTTTAAATTTTATATATATTAATAAATAATTTTATATATAT 16856
OY 2285 ataaacactcacaatccaccaa 2307
DB 16855 AATAATATTTTATTAATAAATAA 16833

RESULT 7
AC007483 AC007483 176174 bp DNA PRI
LOCUS Homo sapiens, clone RP11-72A1, complete sequence. 27-DEC-2000
DEFINITION

```

ACCESSION      AC007483
VERSION        AC007483.7
KEYWORDS       GI:11991320
SOURCE         HTG.
ORGANISM       human.
REFERENCE      Homo sapiens
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE          1 (bases 1 to 176174)
JOURNAL        Birren, B., Linton, L., Nusbaum, C. and Lander, E.
REFERENCE      Homo sapiens, clone RP11-72A1
AUTHORS        2 (bases 1 to 176174)
REFERENCE      Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
AUTHORS        Baker, J., Baldwin, J., Barna, N., Beckert, R., Benn, J., Brown, A.,
AUTHORS        Castle, A., Cerny, D., Colangelo, M., Collins, S., Collymore, A.,
AUTHORS        Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K.,
AUTHORS        Donegan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C.,
AUTHORS        Funke, R., Gage, D., Galagan, J., Garayna, S., Gilbert, D., Grant, G.,
AUTHORS        Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,
AUTHORS        Karatas, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P.,
AUTHORS        Margulis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
AUTHORS        Meldrum, J., Molla, M., Morris, M., Morrow, J., Mychaleckyj, J.,
AUTHORS        Naylor, J., Nioff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
AUTHORS        Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
AUTHORS        Stange-Thomann, N., Stojanovic, N., Stone, C., Sudramanlian, A.,
AUTHORS        Testaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
AUTHORS        Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
TITLE          Direct Submission
JOURNAL        Submitted (05-MAY-1999) Whitehead Institute/MIT Center for Genome
REFERENCE      Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS        3 (bases 1 to 176174)
REFERENCE      Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
AUTHORS        Barna, N., Bastien, V., Boguslavsky, L., Bouhassira, B., Brown, A.,
AUTHORS        Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
AUTHORS        Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
AUTHORS        Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
AUTHORS        Garayna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
AUTHORS        Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
AUTHORS        Jones, C., Karatas, A., Larocque, K., Lamazares, R., Landers, T.,
AUTHORS        Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
AUTHORS        Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
AUTHORS        McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
AUTHORS        Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H.,
AUTHORS        O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
AUTHORS        Phunhahng, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
AUTHORS        Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M.,
AUTHORS        Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
AUTHORS        Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
AUTHORS        Straus, N., Sudramanlian, A., Talamas, J., Testaye, S., Theodore, J.,
AUTHORS        Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A.,
AUTHORS        Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zalnoun, J.,
AUTHORS        Zembek, L., Zimmer, A. and Zody, M.
TITLE          Direct Submission
JOURNAL        Submitted (27-DEC-2000) Whitehead Institute/MIT Center for Genome
REFERENCE      Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS        On Dec 24, 2000 this sequence version replaced gi:11597092.
COMMENT        All repeats were identified using RepeatMasker:
COMMENT        Smit, A.F.A. & Green, P. (1996-1997)
COMMENT        http://ftp.genome.washington.edu/RM/RepeatMasker.html
FEATURES       -----
SOURCE         Location/Qualifiers
SOURCE         1..176174
SOURCE         /organism="Homo sapiens"
SOURCE         /db_xref="taxon:9606"
SOURCE         /clone="RP11-72A1"

```

```

repeat_region /clone_11b="RP11-11 Human Male BAC"
repeat_region complement(2..853)
repeat_region /rpt_family="THE1B-10c"
repeat_region complement(654..1196)
repeat_region /rpt_family="THE1B"
repeat_region 1199..1226
repeat_region /rpt_family="AT_rich"
repeat_region 1237..1269
repeat_region /rpt_family="AT_rich"
repeat_region complement(1448..1661)
repeat_region /rpt_family="LIME"
repeat_region complement(2810..2979)
repeat_region /rpt_family="LIP4"
repeat_region complement(2980..4033)
repeat_region /rpt_family="LIP4"
repeat_region complement(4034..9553)
repeat_region /rpt_family="LIP4"
repeat_region complement(9552..9638)
repeat_region /rpt_family="LIP4"
repeat_region complement(9647..11365)
repeat_region /rpt_family="LIP4"
repeat_region complement(11362..14775)
repeat_region /rpt_family="LIP4"
repeat_region complement(15566..15748)
repeat_region /rpt_family="HAL1b"
repeat_region 15953..16269
repeat_region /rpt_family="MLT1"
repeat_region 17391..17700
repeat_region /rpt_family="ALV45"
repeat_region 17695
repeat_region /note="Poly A region that contains some background noise."
repeat_region 17740..17777
repeat_region /rpt_family="AT_rich"
repeat_region complement(18430..18744)
repeat_region /rpt_family="ALV4"
repeat_region 19230..19286
repeat_region /rpt_family="MLT2"
repeat_region complement(19517..19588)
repeat_region /rpt_family="LIP16"
repeat_region complement(19561..20101)
repeat_region /rpt_family="LIP3"
repeat_region 20102..20125
repeat_region /rpt_family="TTTGn"
repeat_region complement(20126..23539)
repeat_region /rpt_family="LIP3"
repeat_region 23836..23877
repeat_region /rpt_family="AT_rich"
repeat_region 24305..24379
repeat_region /rpt_family="MIR"
repeat_region 24672..24975
repeat_region /rpt_family="MIR"
repeat_region 25026..25163
repeat_region /rpt_family="LRR16c"
repeat_region complement(25738..25806)
repeat_region /rpt_family="MIR"
repeat_region 26037..26162
repeat_region /rpt_family="MERSA"
repeat_region complement(26457..26573)
repeat_region /rpt_family="L2"
repeat_region complement(28095..28212)
repeat_region /rpt_family="L2"
repeat_region 28518..28731
repeat_region /rpt_family="MERS9"
repeat_region complement(29552..29800)
repeat_region /rpt_family="MIR"
repeat_region complement(29996..30092)
repeat_region /rpt_family="LIP3"
repeat_region 30093..30154
repeat_region /rpt_family="L2"
repeat_region 30815..30971
repeat_region /rpt_family="LRR16b"
repeat_region complement(31036..31249)
repeat_region /rpt_family="L2"

```

```

repeat_region 32643. .32716 /rpt_family="AT_rich"
repeat_region complement(32724. .33464)
repeat_region /rpt_family="LIME1"
repeat_region complement(33478. .33643)
repeat_region /rpt_family="FRAM"
repeat_region complement(33646. .34262)
repeat_region /rpt_family="LIME1"
repeat_region complement(34257. .35239)
repeat_region /rpt_family="LIPAL5"
repeat_region 35240. .35307 /rpt_family="(TTTAA)n"
repeat_region complement(35313. .39598)
repeat_region /rpt_family="LIPAL0"
repeat_region 39599. .39901 /rpt_family="AUSg"
repeat_region complement(39902. .41936)
repeat_region /rpt_family="LIPAL0"
repeat_region complement(41945. .42415)
repeat_region /rpt_family="LIPAL5"
repeat_region 42416. .42439 /rpt_family="(TTTG)n"
repeat_region complement(42440. .43554)
repeat_region /rpt_family="LIPAL5"
repeat_region complement(43565. .43805)
repeat_region /rpt_family="LIPAL5"
repeat_region complement(43804. .43967)
repeat_region /rpt_family="LIPAL5"
repeat_region complement(43996. .44296)
repeat_region /rpt_family="AUSc"
repeat_region complement(44299. .45059)
repeat_region /rpt_family="LIM4"
repeat_region complement(45064. .45194)
repeat_region /rpt_family="LIMEc"
repeat_region 45351. .45467 /rpt_family="MIR"
repeat_region complement(45671. .45969)
repeat_region /rpt_family="AUSg"
repeat_region 46325. .46419 /rpt_family="MIR3"
repeat_region complement(46397. .46526)
repeat_region /rpt_family="L2"
repeat_region complement(46565. .46820)
repeat_region /rpt_family="MIR"
repeat_region complement(47841. .48058)
repeat_region /rpt_family="L2"
repeat_region 48305. .48525 /rpt_family="(TTCA)n"
repeat_region complement(49027. .49257)
repeat_region /rpt_family="L2"
repeat_region complement(49283. .49426)
repeat_region /rpt_family="MIR"

```

Query Match	3.1%	Score 77;	DB 86;	Length 176174;
Best Local Similarity	42.9%	Pred. No. 0.0023;		
Matches 819; Conservative	0;	Mismatches 1070;	Indels 20;	Gaps 8;

[illegible]

QY	784	gtttcttaacgggaactctggaagaaagaaataatgtaattcttggaagaacttaaaatag	843
Db	67836	AAATTAATATTTTATCTATTATATATATATATATTTTATATATATATTAATATATATA	678955
QY	844	taaaaggtatcatatggaagaagaagaagaaagaaatagtcgacagaatgga	903
Db	67896	TAATATATATATATAT - ATATTAATATATATATATATATATATATATATATATATAT	67954
QY	904	gaggttcgaggaagaaagcgaatctggagccttgatgctgtgacgcgcgcgtt	963
Db	67955	TATTAAGTATATATTTTATATATATTTTATATATATATATATATATATATATATATAT	68014
QY	964	ttcttaacgcctctcccaactcaaccacatbaaacattctccctattataatt	1023
Db	68015	TTTATGATATAGATATATATTTCTATATATATATATATATTTCTATATATATATATAT	68074
QY	1024	atattcacatgctctatgcttactatgtaaatggcgaccacttaagattacatacat	1083
Db	68075	ATAATTTCTATATACATTTTATATATATATATATATATATATATATATATATATATAT	68150
QY	1084	gtatatactcttaaggatcatcacaaaatgycatgaaacttcgaacttcaactact	1143
Db	68131	GTAATATATATATTTTAT	68187
QY	1144	tgatcatctgaaagtcgtagccttccactgctttggaataatgacgtcgtacctaatt	1203
Db	68188	AAATTTTATATATATATATTTTATATATATATATATATATATATATATATATATATAT	68247
QY	1204	ttaattgacatgctttctgttcgcaacgcttaattctctgatactatgctcaaaattc	1263
Db	68248	AAATTAATATATATATATATATTTTATATATATATATATATATATATATATATATAT	68307
QY	1264	agagcgttcgaagactcttaactaaattcttaaaaaataatcgatlaagagaataagat	1323
Db	68308	TTACATTAAT	68366
QY	1324	tttcacgaacccagtgctgtagtaacgcagcgcgggaatgctctlaaacgacttagctt	1383
Db	68367	TAAATTAAT	68426
QY	1384	ggcttttgatggcttagaactggcatcaggacactccaaacttttgtagctgt	1443
Db	68427	GAT	68486
QY	1444	tgatttaggaatgcgtaaaagacttttattcttaacacagctcgagacttggatcgatat	1503
Db	68487	TTAATAGAT	68546
QY	1504	tacttgaacacttggctggcttcaatgctatgacctatataaacaacacgtaatt	1563
Db	68547	TAAT	68603
QY	1564	ataataggaatttttcggaatttaagcccatcgttaagtataataacatgctgc	1623
Db	68604	ATCTATTAT	68663
QY	1624	gtttcaaatcatcatgatgaagacgtccacgtaagtcgtactactccataacattgcat	1683
Db	68664	TAT	68723
QY	1684	gagagagataatgattataaatttattcttgaagaagaataaagagggaaggttactg	1743
Db	68724	AAATAGAT	68783
QY	1744	ggctgcatgctggaagaaagaagaagaagaaggaacccacacgacccatcatagata	1803
Db	68784	AAATATGAT	68843
QY	1804	tgcagactcttaactcttccctcttattatttcttcaggaacttcttccacttaatt	1863
Db	68844	TAAT	68902
QY	1864	gaaacctccaactatcttaactaataccctccacatgataagaataagaataatataaga	1923


```

Oy 1517 tgggtggttcacgtatcttgccatatacaacaacatcgttaattatatacagattt 1576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 77000 ATATATTTTATTTATATATTTTATATAATATATATTTATTTATATATATAT 76941
Oy 1577 ttccgaatttcaaccatctgtaaglatatataacatgcgtcttccaattca 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76940 ATATATTTTATTTATATATTTTATATAATATATATTTTATATATATATAT 76881
Oy 1637 tatgatgaacgacccacgtatgctactaccctacaatattgcagagagatatgt 1696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76880 ATATTTTATTTATATATTTTATATAATATATTTTATATATTTTATATAATAT 76821
Oy 1697 attataaattttatttgaagaagaataagaggagaggttacttggtggtcagatgt 1756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76820 ATTTTCTTTATATATTTTATATAATATATTTTATATATTTTATACATATATAT 76761
Oy 1757 gaacaacaagaagaagaacggaaccactaagccatcagatatacgacctctat 1816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76760 TTTATTTATCTATTTATATACCCCATATTTTATTTATATATCTTTAATATATATTT 76701
Oy 1817 ctttctccttatttatttcttctcaggaacttttctacttaatacgaacctccaac 1876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76700 TTTATATATATTTTATATAATATATTTTATATATATTTTATATAATATATTT 76641
Oy 1877 tatctaactatacaactccacgttagaataaagaataatataagacattgttgatct 1936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76640 TTTATATATTTTATATAAATATATTTTATATATATTTTATATAAATATATTT 76581
Oy 1937 ttgtaactagaataatattgctcgttaatttctgtaagttcaaatcaaatcttcag 1996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76580 TTTATATATTTTATATAAATATATTTTATATATATTTTATATAAATATATTT 76521
Oy 1997 tagaacaacatattactgcacaaagtagacatatttctgccaacatccagttagc 2056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76520 TTTATATATTTTATATAAATATATTTTATATATATTTTATATAAATATATTT 76461
Oy 2057 tataaggtctagtaaaacaacaacacatctcgtattgccccaaaataaagagag 2116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76460 TATTTATATTTTATATAAATATATTTTATATATATTTTATATAAATATATAT 76401
Oy 2117 aagaatatttcaaaagtgctctctctctctctctctctctctctctctctctct 2176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76400 TATTTATATTTTATATAAATATATTTTATATATATTTTATATAAATATATAT 76341
Oy 2177 agattcaacaagctca 2192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76340 ATTTATATATATTTTA 76325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
AC012492 183250 bp DNA HTG 22-MAR-2001
LOCUS Homo sapiens chromosome UNK clone RP11-224B10, WORKING DRAFT
DEFINITION
SEQUENCE 2 unordered pieces.
AC012492
AC012492.8 GI:13431121
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 183250)
Mammalian, R.H.
The sequence of Homo sapiens clone
unpublished
2 (bases 1 to 183250)
Waterson, R.H.
Direct Submission
Submitted (28-OCT-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Mar 22, 2001 this sequence version replaced gi:12963041.
COMMENT

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0224B10
----- Summary Statistics -----
Sequencing vector: M13; 50%
Sequencing vector: plasmid; 49%
Chemistry: Dye-terminator; 50% of reads
Chemistry: Dye-terminator; Big Dye; 49% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 181369 bases at least Q40
Consensus quality: 182006 bases at least Q30
Consensus quality: 182489 bases at least Q20
Insert size: 175000; agarose-fp
Insert size: 183150; sum-of-coverage
Quality coverage: 7.38 in Q20 bases; agarose-fp
Quality coverage: 7.26 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1187: contig of 1187 bp in length
* 1188 1287: gap of unknown length
* 1288 183250: contig of 181963 bp in length.
Location/Qualifiers
source
1. 183250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="UNK"
/clone="RP11-224B10"
1. 1187
/misc_feature
/feature="assembly_name:Contig1"
1288. 183250
/feature="assembly_name:Contig4"
/feature="end:Spb"
vector_side:left"
BASE COUNT 53820 a 33693 c 33923 g 61705 t 109 others
ORIGIN
Query Match 3.1%; Score 76.4; DB 62; Length 183250;
Best Local Similarity 43.9%; Pred. No. 0.0029;
Matches 585; Conservative 0; Mismatches 726; Indels 23; Gaps 5;
Oy 990 aacacatgaacatctctctctctctctctctctctctctctctctctctctctct 1049
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16477 ACACATATTTTATATTTTATATAATATTTTATATAATATTTTATATAATATTT 16536
Oy 1050 gtaaatgtagcaactaagatattatatacatatgatacatatcttaagtcataca- 1108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16537 ACACATTTTATATTTTATATATTTTATATAATATTTTATATAATATTTTATAG 16596
Oy 1109 -----aaatgltcaagaactttgcaattcaactctactgttcattgtagatcagct 1164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16597 ACATATATTTTATATTTTATATAATATTTTATATAATATTTTATATAATATTT 16656
Oy 1165 ttccacatgttttgaagaataagctcggaatcgtgaactcttcaattagacattgttggtg 1224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16657 TTTCAATATATTTTATATAATATTTTATATAATATTTTATATAATATTTTATAG 16716
Oy 1225 gtaacgttcaattctctgtatattgtagtcaaaatctcagagcgttcagaaactctaca 1284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16717 ATATATTTTATATAATATTTTATATAATATTTTATATAATATTTTATATAATAT 16776
Oy 1285 ctaattcttaaaaaataatcagatta-----agagaataagatttcaatgcacagtgct 1339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db	16777	ATCTTAAATATATATGTTATATTTAAATATATTTAAATATATAGTAAATATTTAAATATATAT	16833
Qy	1340	tgtatagtaacgtagtcgcggaaagtcctaaacgaattatagttcgggttttgatcgtc	1399
Db	16837	TTTAATATATATCTTTATATTTATATATATATATTTAAATATATAGTATATTTTAAATATATTTT	16896
Qy	1400	agaattcgtaactagtaagacccttcaactctttctgtacgtcgtcgtattagtcgta	1459
Db	16897	AATATATATATATTCATTTTAAATATATATATCTTTAAATATATATATTTTATATATATGCTC	16956
Qy	1460	aagagctcttatttcttaaccagctgagacttggatcgatgactgaaacttgg	1519
Db	16957	AATATATATATTCATATATATATTTTTCATATATATATTTTATATATATTTTAAATATATATA	17016
Qy	1520	tgtgttccatgatttggccactatata----aacaaacgtgtaattatatacgaatt	1574
Db	17017	TATTTTATATATATTCATATATATATATATTTTAAATATATATATTTTAAATATATTTTATA	17076
Qy	1575	tttttcggaattttaaaccata-----tcgtaaagtatataacatgcagtcgtttc	1629
Db	17077	TTTAATATATTTTAAATATATATATATTTTATATATATTCATATATTTAAATATATTTTATAT	17136
Qy	1630	aaattcatactgaaacagatccacgtaagtgctactactcctacaatttgcattgagaga	1689
Db	17137	ATATTTATATATTTAATATATCTATATTTTAAATACATATATATTTTAAATCTATATATATATATCA	17196
Qy	1690	gatatgtaattataattttcttggagaagaataagagggaaagcttaacttgggtgga	1749
Db	17197	TATATTTTATATTTTAAATATATATTTTAAATACATATATCTCATATTTTAAATATATTTTAAACA	17256
Qy	1750	tcgaatgtgaaacaaagaagaagaagaagccacacacgaaccatcatatgatatcgaac	1809
Db	17257	TATATTTATATTTTAAATATATATATATTTTAAATACATATATTTTATATATATATATATATAT	17316
Qy	1810	ttctattccttttccctcttatttatttcttctcgaagacttttctcctcaatgaaac	1869
Db	17317	TTAAATATATATATTTTAAATATATATATATTTTATATATATTTTAAATATATATTTTATATTTTA	17376
Qy	1870	tccaaacatctaaactaaatacaccoccatgtagaaataagaanaattatataagattcgt	1929
Db	17377	TATATCTTTTATATTTTAAACATATATATTTTATATATATTTTAAATATATATTTTATATTTT	17436
Qy	1930	tgaatttttgtaactagaaaatatatttgcctcgttaatttctcgtgaagttaacacat	1989
Db	17437	ATATTTTATATATATATATATATATATTTTAAATATATGTTATATATTTTATATTTTAAATAA	17496
Qy	1990	ttttcagagaagaacaaattactcgcgaagaagtgatcatatttcttgcacaacatc	2049
Db	17497	TATATTTTAAATATATATTTTATATTTTAAATATATTTTAAATATATTTTAAATATATTTTATAT	17556
Qy	2050	agttgcatatagggtctgtagtaaaaaaacaacacacatctcgtttgcgcccaaaaaataaa	2109
Db	17557	TTATATATTTTAAATATATTTTAAATATATATATTTTATATATTTTAAATATATATTTTAA	17616
Qy	2110	gagagagaagaatatctgtccaagaagtgctcctctcctcctaattatgcttccactaa	2169
Db	17617	TATATATTTTAAATTTTATATATTTTAAATATATATTTTAAATATATATTTTAAATATATATA	17676
Qy	2170	cccaattagatccaacagcttcaacaagtcacaa---agataaacatgagacaacaattc	2226
Db	17677	TTTTTATTTTATATATATATTTTATATTTTAAATATATATTTTATATTTTATATATATATA	17736
Qy	2227	gatgcaaaaaaatcctcttttcaatgctcttttttattctcctagctctttaaattactaat	2286
Db	17737	TACATTTTAAATATATATTTTATATTTTAAATATATATTTTAAATGCAATATATATATATATAT	17796
Qy	2287	aaaaactcacaaat 2300	
Db	17797	ATAATGTCGAATAT 17810	

LOCUS	CNS01JRG	879 bp	DNA	STS	17-FEB-2000
DEFINITION	Anopheles gambiae STS T7 end of clone 14D07 of NotreDame1 library from strain PEST of Anopheles gambiae (African malaria mosquito), sequence tagged site.				
ACCESSION	AL147405				
VERSION	AL147405.1 GI:7005551				
KEYWORDS	STS.				
SOURCE	African malaria mosquito.				
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.				
REFERENCE	1 (bases 1 to 879) Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
JOURNAL	- Web : www.genoscope.cns.fr) 2 (bases 1 to 879)				
REFERENCE	Roch,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J. Direct Submission				
AUTHORS	Submitted (16-FEB-2000) BMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France				
JOURNAL	This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.				
COMMENT	Location/Qualifiers				
FEATURES	1..879 /organism="Anopheles gambiae" /strain="PEST" /db_xref="taxon:7165" /clone_14D07 /clone_11b="Notredame1" /note="end : T7"				
BASE COUNT	66 a 66 a 38 c 1 g 670 t 104 others				
ORIGIN					
Query Match	3.0%; Score 75; DB 53; Length 879;				
Best Local Similarity	40.7%; Pred.No. 0.0071;				
Matches 201; Conservative	51; Mismatches 241; Indels 1; Gaps 1;				
433	agacaaacagagaaatcatgatgagccatgcatctaagtgcgctgtgttaattagaggtc	492			
787	AAAAAAAGGAAAAAAGTAAATAAAAAATATGTGAACAKAAAAAATATAGT	728			
493	cgcatacatataaacagtagagacatagataaataatgaaacacacacca-aaagagtc	551			
727	AAAAAARATAGCAAADKAKWAAAAAAAKTAAARGWAAAAAATAAAAAAATAAAAAA	668			
552	ggaacatcctaagctgttagagagaataatagtcctcaggtcggaagatccaagaagagac	611			
667	GGAATTAATAAAAAAAMWAAAAAAGAAAAAAGRTTATKMKAAKAWGAAAAAAKAGAA	608			
612	atgaagaggtatataagctctaaacaaatgagcagctagtcggaagaggtcttaact	671			
607	AATTWGWGCTDAAAAAAYAGAAAAAAGAGAGKAGKGAARAAAGAAKAAKAAKAA	548			
672	tgaacaaagtagatctgtagaagaacaagaagaagatgcgcctagttctagata	731			
547	AAAAAAATGTTTARAATAAGAAAAAATAKATRTTAAAAAAMWAKAGAAAKRGWAGA	488			
732	ataattacacattcgtgttataataagtaagagaatatgacacattgtgttctcta	791			
487	AAARAGATWAKAKAAAAAATAAAAAATDAAAAATATWMTWBAAAAAAAGAAKAAKAT	428			
792	cgggtgaatctgttagaagaaaaaataagtaaatcttctgagaaatctaaatagtaagag	851			
427	TCDRAAAAAAATAAAAAAANANAAAAAATAAAAAAATAAAAAAATAAAAAAATA	368			
852	tatatctggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaggttag	911			

	Query Match	2.9%; Score 72.8; DB 86;	Length 176174;
	Best Local Similarity	42.2%;	Pred. No. 0.011;
	Matches 718; Conservative	0; Mismatches 967;	Indels 15; Gaps 5
OY	gttgcacatcacaataaacagtagacatatgagataaatatgaacacacaccacaaagaag	549	
Dd	69501 GTGCTATTATTTTATTTTAAATGTATAAATAATTAATTAATTTTTTTTATAAATAAT	69442	
OY	tgggaatactcaataagtatgtcagaagaataaagtcctcaggttggagatticaagaagagag	609	

Db	69441	TAATATATTCCTTATATATATATAATTAATTTATATATATGATATATATATATATTTATATTT	69382
Qy	610	acaaatgaagggtctatagactctaaacaaaatgcatgacttagtggagaaggctctta	668
Db	69381	ATATATAATATATGCAAAATTTTATATATATATAAATTTTATAAATATATATATATATACATA	69322
Qy	669	aattgaacaagtaggattcgaagcaacgaagaaacaaagaacgacgacctagattctcgag	728
Db	69321	AATTTTATATATATAATATATATATTTATATATATATATATATATATATATATATATATATAGAT	69265
Qy	729	ataataatcacattgcctgttatataaggaagagatctgacattgcttgcttc	788
Db	69261	ATATATATATATATATATAAATTTATATATAGATATATTTATATATATATATATAATATATATA	69202
Qy	789	ttacaggctaaatctgagagaaaaaaatagatatacttggagaaaactaaatagataag	848
Db	69201	AGTATATATATAATATATAAATATATAAATTTATATATTTATATATATATATATATATAATA	69142
Qy	849	aggtatatactgagagaaagagagaaagggaaaaatagctgacagagatcggagagaggt	908
Db	69141	ATTA	69082
Qy	909	tagagagcgaagggcaaatgctgagactcttgatagatctgcatgacgcgctcaagcttctt	968
Db	69081	GAGAAATATATAAATTTTATATATATTTCTATATATATATATATATATATATATATATATATAT	69022
Qy	969	caagcctgtcccaaccacactcaccctcgaacatctccctcatttbaattatatt	1028
Db	69021	AATATATAATATAGTTATATATATATATAATATACAAATATATAAATATATAATATATATA	68962
Qy	1029	cacatgctcatgctactatgctaaatggtgacccacttaagtatctatcatcatgata	1088
Db	68961	ATTATATATATATATATATATATATAAATGCTATATATTTGCTATACAAAATATATATTTTA	68902
Qy	1089	tattctataggtctcatcaaaaaatggatcgaacatttgcatttcaatctactctgttc	1148
Db	68901	TATATAAATATATACATATATATATATATATAAATAATATAAATATATA-----TTTATA	68848
Qy	1149	attgtagatgctagctcttccacatgcttctgaaaattagctcgatctcgaaatcttcaat	1208
Db	68847	TTTATATAAAATTTATATTTTATATTTTATATTTTATATTTTATAAATTTATATATTTTACA	68786
Qy	1209	tagcatgttttgctgctgaacgcttaattcttcgattatgactgctcaaaaatctcgagc	1268
Db	68787	TATATATATATTTATATATATATATATATATATTTATATATATATATATAATATATATATAC	68728
Qy	1269	gttcggaactctcacactaatcttcttaaaaaaatcgatcgaagaagaatagagcttcaa	1328
Db	68727	TATATATATAAATAGTA	68668
Qy	1329	tgcaccagctgtgatagtagtagtcgcggaatgctcgaagaacgataatgagttgggtg	1388
Db	68667	TATATATATTTTATATATATATATATAATAGATATATATATATATATATATATATATATATAT	68608
Qy	1389	tttgattggttagaatctgctatagtaggaacacttcaactcttttgcttagcttgctgatt	1448
Db	68607	AGAT	68548
Qy	1449	tagagatgctgaagagctcttcttaattcttaaccaggttgagactgggactcgatagactt	1508
Db	68548	TATATATATATTTATATATATATATATAGATATATATATATATATATATATATATATATAT	68489
Qy	1509	gaaacactgggttgcttcatgctatcttggtcctatatataaacaacacgtataatata	1568
Db	68488	AAATATATTTATATATAAATA	68429
Qy	1569	cggattctttctcgaaattctaacccatctctgaagtatataaacaatgactgctttt	1628
Db	68428	TCTATATATATATTTAT	68369
Qy	1629	caaatctatagtgaaagcatccacgtaagtgctactactcctcaaatatgcatgagag	1688

D	68368	TATATTTTTAAATAAATATATATTAATTAATTAATTAATTAATTAATTAATGA				68309
Oy	1669	agataatgattacaatatcttgcgaagaacaaataaggggaggtcacttggtcg				1748
D	68308	ATATACATATATATTTATTCATATAATTTTTATATATAAAAATATATATATATTT				68253
Oy	1749	atgatgtgaacaacaaaagaagaagaacgcgaaccctaagccattcatgatcgcac				1808
D	68254	ATATATTTAAAATTCATATATAAATTTATATTTATAAATTTTATATATAATTTTATAT				68195
Oy	1809	ccticatcttccttccttatcttaattttctcgaacctttctcaattaatgaaac				1868
D	68194	TAAATTTATATATTTTATATATTAATAATTAATATATATTTTATATTAATAATATAT				68135
Oy	1869	ctccaactatcatactaatcacctcccagttagaataagaacattatagaatctg				1928
D	68134	ATACAAAATTTATATGTCTTAAAAATTTATATATATATATATATATGATATAGAAATAT				68075
Oy	1929	tgtgatcttctgtaactagaaa-ataatcttgtcctgttaattttcgttaagttaa				1987
D	68074	TATTTAAATTTATATATGCAATAATATATATATATATATAGAAATATATATCCTTATACAT				68015
Oy	1988	attcttcagtagaacacaataattactgcaaaaaagtagatcaltaattcttgc				2047
D	68014	TATATATGTTTATATTTTAAATGCATTTATATATATAAATATATATATATATATCTATAT				67955
Oy	2048	tcagttagctcagggtctgttagcaaaaacacacatcttgattgcccaaaaaa				2107
D	67954	TATTTATCAAATATCGTAGTATTTTATATATATATATATATATATATATATATATATAT				67895
Oy	2108	aaggagagggaagaatattgctcaaaaagsgctctctccctcatatagtcttac				2167
D	67894	ATTATATTTTATATATTTATATATATAATATATATATATATATATATATATATCATTAAT				67835
Oy	2168	aaccaattagattcaaaca	2187			
D	67834	TAGAATTTTATATTTTAAATA	67815			
RESULT 12						
AF203688/C						
LOCUS	AF203688	5665 bp	DNA	PLN	02-JAN-2001	
DEFINITION	Canavalia lineata ornithine carbamoyltransferase OOCtl (OOCtl)					
ACCESSION	AF203688					
VERSION	AF203688					
KEYWORDS	AF203688.1 GI:12003187					
SOURCE						
ORGANISM	Canavalia lineata.					
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;					
	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;					
	Fabales; Fabaceae; Papilionoideae; Canavalia.					
REFERENCE	Yoo,S.K., Lee,Y. and Kwon,Y.M. Properties of ornithine carbamoylases purified from Canavalia lineata leaves and their cDNA cloning					
AUTHORS	Unpublished					
TITLE	2 (bases 1 to 5665) Yoo,S.K., Lee,Y. and Kwon,Y.M.					
JOURNAL	Direct Submission					
REMARKS	Submitted (10-NOV-1999) Biology, Seoul National University, San					
JOURNL	56-1, Shinrim-Dong, Kwanaek-Gu, Seoul 151-742, Korea					
FEATURES	Location/Qualifiers					
source	1..5665 /organism="Canavalia lineata" /db_xref="taxon:28957" join(1702..1867,2204..2538,4005..4136,4240..4515, 5276..5665) /gene="OOCtl" /product="ornithine carbamoyltransferase OOCtl" 1702..5665 /note="OOCtl" join(1714..1867,2204..2538,4005..4136,4240..4515,					
CDS						

Query Match	Best Local Similarity	Matches	Score	DB	Length
419: Conservative	44.3%	0: Mismatches	71.8	13	5665
419: Indels	0.02%	522: Indels		4	Gaps
419: Gaps					
1540	1896	769	878	g	2122
1050	1896	769	878	g	2122
1600	1896	769	878	g	2122
990	1896	769	878	g	2122
1660	1896	769	878	g	2122
911	1896	769	878	g	2122
1720	1896	769	878	g	2122
871	1896	769	878	g	2122
1780	1896	769	878	g	2122
811	1896	769	878	g	2122
1840	1896	769	878	g	2122
751	1896	769	878	g	2122
1900	1896	769	878	g	2122
691	1896	769	878	g	2122
1960	1896	769	878	g	2122
631	1896	769	878	g	2122
631	1896	769	878	g	2122
2020	1896	769	878	g	2122
571	1896	769	878	g	2122
2080	1896	769	878	g	2122
512	1896	769	878	g	2122
2140	1896	769	878	g	2122
452	1896	769	878	g	2122
2198	1896	769	878	g	2122
392	1896	769	878	g	2122
2258	1896	769	878	g	2122

Search completed: November 22, 2001, 22:16:10
Job time: 31672 sec

Search completed: November 22, 2001, 22:16:10
Job time: 31672 sec

Gene	Accession	Position	Feature	Annotation
AC012492	3,1	183250	62	AC012492 Homo sapi
CNS01JRG	3,0	879	53	AL147405 Anopheles
AC007483	2,9	176174	86	AC007483 Homo sapi
AF203688	2,9	5665	13	AF203688 Canavalia
AC005505	2,9	192929	60	AC005505 Plasmid
AL12783	2,9	185800	81	AL12783 Homo sapi
S96842	2,8	2305	96	S96842 SP96-spore-
AF211124	2,8	7347	2	AF211124 Carsonella
AC010727	2,8	148061	87	AC010727 Homo sapi
AC011718	2,8	94384	87	AC011718 Homo sapi
AC019041	2,8	183358	89	AC019041 Homo sapi
CEY683B	2,7	69709	6	CEY683B
AL022655	2,7	110000	83	AL022655 Caenorhab
CEY683	2,7	25356	83	CEY683
CEY87G2	2,7	330612	83	CEY87G2
AL022597	2,7	160401	87	AL022597 Caenorhab
AC009277	2,7	253305	96	AC009277 Homo sapi
PEMAL3P7	2,7	184535	86	PEMAL3P7
CNS05TCJ	2,7	208546	76	AL355100 Homo sapi
AC083827	2,7	1141	10	AC083827 Homo sapi
AX083744	2,7	231912	77	AX083744 Sequence
AC087566	2,7	19517	6	AC087566 Mus muscu
DMU37541	2,7	165097	84	DMU37541
CNS05TCR	2,7	964	54	CNS05TCR
CNS07EBR	2,7	253305	96	CNS07EBR
PEMAL3P7	2,7	1141	10	PEMAL3P7
AX083744	2,7	186666	76	AX083744 Sequence
AC079409	2,6	156060	61	AC079409 Homo sapi
AC004153	2,6	174384	61	AC004153 Plasmid
AC009524	2,6	4601	6	AC009524 Homo sapi
DMU11584	2,6	136098	86	DMU11584
AC006970	2,6	6572	6	AC006970 Homo sapi
DDICNPA	2,6	83440	68	DDICNPA
AC024285	2,6	185603	74	AC024285 Homo sapi
AC071336	2,6	170157	79	AC071336 Homo sapi
AL355372	2,6	174565	74	AL355372 Homo sapi
AC069525	2,6	178087	85	AC069525 Homo sapi
AC005089	2,6	178087	85	AC005089 Homo sapi
foranism="Arabidopsis thaliana"				
/variety="Columbia"				
/db_xref="taxon:3702"				
/chromosome="4"				
2053..3409				
/gene="F19F18.10"				
join(2053..2268,2351..2551,2625..2790,3003..3409)				
/gene="F19F18.10"				
/EC_number="1.1.1.7"				
/note="Contains Peroxidases signatures [DM1ALSGAHTI]				
contains EST g:576577, Z35618, H37596, R90036, Z35219,				
AA713329, AA721954"				
/codon_start=1				
/product="peroxidase, prxr2"				
/protein_id="CA38291.1"				
/db_xref="GI:4468977"				
/translation="MYVYKNTMLILSLICITLDSNQLRNFYAGSCNRYQIVR				
NAVOKVQTFPTTPTATLELYHDFPVNGCDASVIASTINNKAEKEHEMLSLAGOG				
EPQVIRAKALDAVPRGRKRSACDILMTATDVVNLGAPDYVEIGRLDLSSTAA				
SYGGKLPHTDVKLTSLIFANGLSLMDIALSCAHLGFACTEYRNRYLTENKTT				
KVDPYVNRKDYVELKASCPNRIDPVAIMNPETPRORDNYRYNKLQGRGLFSDVQ				
LEFDRSKETPEVLMANNGLQFNAFINSMIKRGVGTSGNSNIRDCGAFN"				
2053..2268				
/gene="F19F18.10"				
/number=1				
2269..2350				
/gene="F19F18.10"				
/number=1				
2351..2551				
/gene="F19F18.10"				
/number=2				
2552..2624				
/gene="F19F18.10"				
/number=2				
2625..2790				
/gene="F19F18.10"				

ALIGNMENTS

FEATURES	Source
LOCUS	ATF19F18
DEFINITION	Arabidopsis thaliana DNA chromosome 4, BAC clone F19F18 (ESSA project).
ACCESSION	AL035605
VERSION	AL035605.1
KEYWORDS	GI:4468976
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 91740) Beyan, M., Koetter, P., Hempel, S., Ertan, K.-D., Bancroft, I., Mewes, H.W., Mayer, K.F.X. and Schneeller, C.
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 91740) EU Arabidopsis sequencing, project.
AUTHORS	Direct Submission
JOURNAL	Submitted (03-MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopfersplitz 18a, D-82152 Martinsried, FRG, E-mail: schuellemaps.biochem.mpg.de, mayermips.biochem.mpg.de Project Coordinator: Mike Beyan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.beyan@bbsrc.ac.uk
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosome 4 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ .
FEATURES	Location/Qualifiers
Source	1..91740

/gene="F19F18.20"
 /number=3
 5893..6024
 /gene="F19F18.20"
 /number=3
 6025..6431
 /gene="F19F18.20"
 /number=4
 complement(join(9998..10536,10675..10858))
 /gene="F19F18.30"
 complement(9998..10536)
 /gene="F19F18.30"
 /number=1
 9998..10858
 /gene="F19F18.30"
 complement(join(9998..10536,10675..10858))
 /gene="F19F18.30"
 /note="similarity to predicted protein, Arabidopsis thaliana contains EST gb:R65200"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB38293.1"
 /db_xref="GI:4468979"
 /translation="MSCNCRVLRKGCSECTILRPLCLQNIIESAQHATYFVAKFEG
 RAGLSFISSVPELQRPALFQSLFEACGRTPNPNVGAAGMLMTNRMHQAAYETVL
 NGRTLRPIISDLLESPSLMISCDSESSIMHWDVSRQTHRCFRSTSRSTEMQASLVNR
 RLKSDSDLDLVNHGLTLTAFAVPVPLPSPSFCVVKGRDPSPESESVTSCWEN
 GMRGNKOKRNGEKKLNTLFY"
 complement(10537..10674)
 /number=1
 complement(10675..10858)
 /gene="F19F18.30"
 /number=2
 14033..14251
 /gene="F19F18.40"
 /number=1
 14033..16078
 /gene="F19F18.40"
 join(14033..14251,14346..14583,14704..14834,14990..15225,
 15334..15577,15691..15798,15956..16078)
 /gene="F19F18.40"
 /BC_number="3.5.1.49"
 /note="strong similarity to formamidase, Methylophilus methylotrophus, Pir2:S74213
 contains EST gb:T6402, AA404801, T22522"
 /codon_start=1
 /product="formamidase-like protein"
 /protein_id="CAB38294.1"
 /db_xref="GI:4468980"
 /translation="MAPLSRPRVVVVDLKKKPMQODPLHNMHPPIPPVAEYKAGEE
 FVEMIDAMGVIKNDSDIKKNVLTTHHSGIRYVDEGVAAKAGDILAVTIC
 NIGPLPDEMGKTFGSDRENGGFLTDHPKCAKAIWIEGITYASPOIPGRFPELT
 HGCVIGTAPSNELRLIWDREKQLESGVESLTLCEIEGTEPWEMLAEAAITLGR
 EENGNDIKNLSGSKYLPVVEGANLSTGDMHFGQSGEISFCALMSFLLEKCK
 EIRRMGOEYILPMGPPLHVNIFIEIGVPEPFSMLVEGTSVDESKOHLIDATV
 AYKRAVLNADILFKEGYSKEQVYVILLSCPCGRSGTVDSPNVAITAIATPAIFDO
 DIRPKRKYPVGRPRVRAKFDVLSKSTTDGKLPTTKNSSSS"
 14252..14345
 /gene="F19F18.40"
 /number=1
 14346..14583
 /gene="F19F18.40"
 /number=2
 14584..14703
 /gene="F19F18.40"
 /number=2
 14704..14834
 /gene="F19F18.40"
 /number=3
 14835..14989
 /gene="F19F18.40"
 /number=3

exon 14990..15225
 /gene="F19F18.40"
 /number=4
 15226..15333
 /gene="F19F18.40"
 /number=4
 15334..15577
 /gene="F19F18.40"
 /number=5
 15578..15690
 /gene="F19F18.40"
 /number=5
 15691..15798
 /gene="F19F18.40"
 /number=6
 15799..15955
 /gene="F19F18.40"
 /number=6
 15956..16078
 /gene="F19F18.40"
 /number=7
 16079..16388
 /gene="F19F18.50"
 /number=1
 16389..16806

Query Match 100.0%; Score 2502; DB 13; Length 91740;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaagcagagcgctggggcttctcctaataatgtagaagaactgcatgagaacattc 60
 DB 59721 AGAAGCAGAGCGCTGGGGCTTCTCTTAATATGTGAAAGAACTGATCATGAGACATT 59780
 QY 61 gactaccagagatggtgtagctcataagatgataatctcactgcatgactgctgac 120
 DB 59781 GACTACCAAGATGGTGTGACTCTAAGATGTAATATCTACTGCAATATGCTTACGC 59840
 QY 121 taggcataatgtagattgtagcaacttctcctaataatgtagttggaatttagcatgat 180
 DB 55841 TAGGCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 55900
 QY 181 atagcatatctaaatgtagcgaacttctcctaataatgtagaagaatgtagatgac 240
 DB 59901 ATAGCATATATCTAATATGTCGGAACCTTCTCATATGATGATGATGATGATGAT 59960
 QY 241 gtaatgtagttgctgtgtaataataaataaataaataaataaataaataaataaataa 300
 DB 59961 GTAATGTAGGTTGCTGTTGTAATATCAAAATACATCATCTTATGTTTATGATTT 60020
 QY 301 tatttatttttaataatgtagtctgtagcgtgtagcgaataaatttccaatttgagac 360
 DB 60021 TATTTATTTTATATATGCTGCTACGTAGCGCATTAATATTCATATTTGAGAC 60080
 QY 361 ttccggaattttaaacaataaataaataaataaataaataaataaataaataaataaataa 420
 DB 60081 TTCCGGATTTTAAACAAATTAACAAATGCGCATGCGTCCGGGGGATAGACAAATT 60140
 QY 421 tgccttgtagcgaacaacgagaataatcatgtagcgcctatgcatgattgagtcggt 480
 DB 60141 TGCCTTGATCGAGCAACGAGAAATATCATGATGATGATGATGATGATGATGATGAT 60200
 QY 481 taattgaggttcgcataataaaccgtagacataatgtagaataatgtagaacaaca 540
 DB 60201 TAATTGAGGTTCGCATATACATTAACATGACATGATGATGATGATGATGATGATGAT 60260
 QY 541 ccaaaaaagtggaataatcaataatgtagaagaataaataaataaataaataaataaataa 600
 DB 60261 CCAAAAAAGTGGGAAATCTTAATTAAGTGTAGCAATTAATTAAGTGTAGCAATTAAT 60320
 QY 601 aaagagaggaacaatgaagggtatagactctaaacaaataatgtagaactgtagtga 660
 DB 60321 AAAGAGAGGACATGAAGGATATATGACTCTAAACAAAAATGGCATGACTTATGAGGACA 60380

OY 661 gggcttaattgaacaagtagattgagacagaagaacaaagacatgcccctaga 720
 DB 60381 GGGTTTAAATTGAAACAAGTAGATGAGAAACAAGAAACAAGAACATGCCCCCTAGA 60440
 OY 721 ttctcgagataaataacacatgctgttataataagtaagaataatgacacattg 780
 DB 60441 TTTCGAGATATATATACACATGCTGTTTATATAGCTAAGAGATATGACACATGCG 60500
 OY 781 ttggtttcttaagggtaaatgctgaagaagaaataatgataattgagaataatctaa 840
 DB 60501 TTGGGTTCTTACGGGTAATGGAAGAAAAAAATATATATTTGACAAAAATCTAAAA 60560
 OY 841 tagtaaaagagcatatattgagaagaagaagaagaagaagaagaagaagaagaagaatg 900
 DB 60561 TAGTAAAGGATATATATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 60620
 OY 901 agagaaggttaaggagcaaaagcaaatgtgagacttgaatgattgacagccgctcag 960
 DB 60621 AGAGAGGTTAGGAGGCAAGGCAAAATGTGAGCTTGATGATGATGATGATGATGATGATG 60680
 OY 961 cttcttcaagcctgctccac 1020
 DB 60681 CTTTCTTCCAGCCGCTCCACTCCACTCCACTCCACTCCACTCCACTCCACTCCACTCC 60740
 OY 1021 attataatcactgtctcattgttactatgtaaatgtgacacacacacacacacacac 1080
 DB 60741 ATTATATATACATGCTCTATGTTACTATGTTAAATGTTAGTAAATGTTAAATGTTAA 60800
 OY 1081 catgtatatacttataagtatcatatacaaaatgctcatgaacttttgcaatttcaatc 1140
 DB 60801 CATGTATATATCTTATAGGATATCATATAAATGTCATGAATTTTGCAATTTTCAATCT 60860
 OY 1141 actgttcatctgtagatgctgacgtcttcaacatgttgaagaatgactgagatcgaat 1200
 DB 60861 ACTGTCTCATGAGATGCTAGCTCTTCAATGTTTGAATTAATGCTGATCTGATTAAT 60920
 OY 1201 tcttaataagactgtttgt 1260
 DB 60921 TCTTAAATTAAGCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 60980
 OY 1261 ttcgaagcgttcgaagaactcttacaacttcttcaaaataacacacacacacacacac 1320
 DB 60981 TTGAGAGCGTTGAGAACTCTTACATATTTCTTAAATAATATGATTAAGAGAAAAATAG 61040
 OY 1321 agtttcatgacacagctgttgaatgaatgtagtcgagaaatgctcaaaacacacacac 1380
 DB 61041 AGTTTCAATGACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 61100
 OY 1381 ttgggtgtttgt 1440
 DB 61101 TTTGGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 61160
 OY 1441 tgttgaattgagatgagcgaagaagccttttatttaccacacacacacacacacacac 1500
 DB 61161 TGTGATTTAGGATGCGTAAAGAGTCTTTTATTTTACACAGTTGAGATTTGGGATCGA 61220
 OY 1501 taagtactgaacacactgt 1560
 DB 61221 TAGTACTGTAACACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 61280
 OY 1561 attataagagatttttttccggaatttcaagcacaatctgtaagatatataacatgcat 1620
 DB 61281 ATTATATACGAGATTTTGTGGAATTTTACGCCATATCTTAAATATATATATATATATAT 61340
 OY 1621 gtgcgtttcaaatcatatgataagacacacacacacacacacacacacacacacacac 1680
 DB 61341 GTGCGTTTCAAAATTCATATGATGATGATGATGATGATGATGATGATGATGATGATG 61400
 OY 1681 catgagagagatatgtattataaatttatcttgaagaagaagaagaagaagaagaagttac 1740
 DB 61401 CATGAGAGAGATATGATATTTATTTATTTTGAAGAAGAAATATAGAGAGAGGTTAC 61460

OY 1741 ttgggtgagatcgatgtgaaacaaagaagaagaagaagaagaagaagaagaagaagaaga 1800
 DB 61461 TTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 61520
 OY 1801 atatcgaccttcttactcttctctctctctctctctctctctctctctctctctctct 1860
 DB 61521 ATATCGACCTTCTTATCTTTTCCCTTTTATTTTATTTTCTCGAGCATTTTCTACTT 61580
 OY 1861 aatgaacctccaaactatccaaactcaactcaactcaactcaactcaactcaactcaact 1920
 DB 61581 AATGAACCTCCAACTATCTTACTTATACCTCCCACTGATGATGATGATGATGATGATG 61640
 OY 1921 agatatgtgatatttgaactgaagaataatatttgcctgtaatttttcgtaagtta 1980
 DB 61641 AGATATGTTGATTTATTTGTAAGTAAATATATTTTCTGTAATTTTCTGTAAGTTA 61700
 OY 1981 aatcaacatttttcagtagaagaacaaatattactgcaaaagtaggaatcatttttgc 2040
 DB 61701 AATCAACATTTTTCAGTGAACAAATATTTACTGCAAAAGTAGATCATTTATTTTGTG 61760
 OY 2041 caaaatccagtttagctataggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2100
 DB 61761 CAAAATCTCAGTTACTATGAGGTTGTAGTAAACAAACAAACATCTTGAATTTGCCCA 61820
 OY 2101 aaaaataagagagagaagaataattgttcaaaagtgctctctctctctctctctct 2160
 DB 61821 AAAAATTAAG 61880
 OY 2161 ttcactaaacccaatttagattcaaaagcttcaaaagcttcaaaagcttcaaaagcttcaaa 2220
 DB 61881 TTCACTAACCCCAATTAGATTAACAAAGCTCAAAAGCTCAAAAGCTCAAAAGCTCAAA 61940
 OY 2221 caattcgatgcaaaaaatccctcttcttcaatgctcttcttctctctctctctctct 2280
 DB 61941 CAATTGAGAGCAAAAAATCCCTTTTCAATGCTTTTATTTTATTTTCTGATCTTAAAT 62000
 OY 2281 actaataaaactccaactcaactcaactcaactcaactcaactcaactcaactcaactca 2340
 DB 62001 ACTAATTAATACTCACAATTCACCAACCAATTTCTTCAACCTCACTTCAATGATTT 62060
 OY 2341 taccacatccacacgagaaacacaaagaagaagaagaagaagaagaagaagaagaagaaga 2400
 DB 62061 TACCCATCCCAACCGGAGAAACAAAGAAAAAAATATATATATATATATATATATAT 62120
 OY 2401 cacatgatgctgatgataatatacaacaaagatataatctttagatattgtgtgtctc 2460
 DB 62121 CACATGATGCTGATGCAATATACACAAACAAAGATATATATATATATATATATAT 62180
 OY 2461 cttctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2502
 DB 62181 CTTTCTTCTATTCATTTTCTTATTCATTAATAAAAAAATAG 62222

RESULT 2
 ANCHOR187
 LOCUS ATCHR187 196339 bp DNA PLN 16-MAR-2000
 DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 87.
 ACCESSION AL161591
 VERSION AL161591.2 GI:7270703
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicot; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1. (bases 42610 to 143618; 123423 to 196339)
 Rose, M., Hempel, S., Entian, K.-D., Mewes, H.-W., Lemcke, K. and
 Mayer, K.F.X.
 JOURNAL Unpublished
 REFERENCE 2. (bases 1 to 196339)
 AUTHORS EU Arabidopsis sequencing project.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer